

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 14:54:57 ; Search time 36.6716 Seconds  
(without alignments)  
69.343 Million cell updates/sec

Title: US-09-641-801-31  
Perfect score: 48  
Sequence: 1 VESYVPLFP 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	48	100.0	9	2	AAW42101	Aaw42101 Colostrin
2	48	100.0	9	4	AAB72529	Aab72529 Colostrin
3	48	100.0	9	4	AAB72276	Aab72276 Colostrin
4	48	100.0	9	4	AAB72561	Aab72561 Colostrin
5	48	100.0	9	5	AAO14607	Aao14607 Neural ce
6	48	100.0	9	5	AAM51067	Aam51067 Colostrin
7	48	100.0	9	5	AAE20258	Aae20258 Colostrin
8	44	91.7	8	4	AAE07189	Aae07189 Colostrin
9	44	91.7	9	4	AAE07199	Aae07199 Modified

10	38	79.2	249	3	AAB18662	Aab18662	A human r
11	38	79.2	249	3	AAy78810	Aay78810	Hydrophob
12	38	79.2	249	4	AAM39162	Aam39162	Human pol
13	38	79.2	288	3	AAB42118	Aab42118	Human ORF
14	38	79.2	288	4	AAG81301	Aag81301	Human AFP
15	38	79.2	326	4	AAM40948	Aam40948	Human pol
16	36	75.0	206	4	AAB59497	Aab59497	Human sec
17	36	75.0	212	4	AAB59515	Aab59515	Human sec
18	36	75.0	254	5	AAU75790	Aau75790	Human pro
19	36	75.0	254	6	ABG99915	Abg99915	Human nov
20	36	75.0	318	4	ABB12293	Abb12293	Human PRO
21	36	75.0	393	5	ABB80589	Abb80589	Human sbg
22	36	75.0	452	6	ABU44622	Abu44622	Protein e
23	36	75.0	523	4	AAE02629	Aae02629	Human sec
24	36	75.0	523	5	AAE16939	Aae16939	Human UDP
25	36	75.0	523	5	ABB80590	Abb80590	Human sbg
26	36	75.0	523	5	AAE15435	Aae15435	Human dru
27	36	75.0	523	5	ABG31867	Abg31867	Human dru
28	36	75.0	523	6	ADA55359	Ada55359	Human pro
29	36	75.0	526	4	ABG20645	Abg20645	Novel hum
30	35	72.9	97	2	AAy34653	Aay34653	C. pneumo
31	35	72.9	152	4	AAG91462	Aag91462	C glutami
32	35	72.9	305	7	ADB64516	Adb64516	Human pro
33	35	72.9	458	6	ABU29136	Abu29136	Protein e
34	35	72.9	466	5	ABB55048	Abb55048	Lactococc
35	35	72.9	471	6	ABP57482	Abp57482	Mycobacte
36	35	72.9	548	2	AAy37206	Aay37206	Protein i
37	35	72.9	614	6	ADA35374	Ada35374	Acinetoba
38	35	72.9	1040	2	AAW97216	Aaw97216	An isoleu
39	35	72.9	1041	3	AAy56782	Aay56782	C. tracho
40	35	72.9	1041	7	ADC34773	Adc34773	Chlamydia
41	35	72.9	4134	2	AAy31946	Aay31946	Plasmodiu
42	34	70.8	12	7	ADE12722	Ade12722	Hev b III
43	34	70.8	12	7	ADE12457	Ade12457	Hev b III
44	34	70.8	64	4	AAM91969	Aam91969	Human dig
45	34	70.8	64	4	AAU22497	Aau22497	Novel hum

# ALIGNMENTS

## RESULT 1

AAW42101

ID AAW42101 standard; peptide; 9 AA.

XX

AC AAW42101;

XX

DT 09-SEP-1998 (first entry)

XX

DE Colostrinin derived nanopeptide.

XX

KW Colostrinin; nanopeptide; NP; central nervous system; CNS;

KW neurological disorder; mental disorder; dementia; Alzheimer's disease;

KW motor neurone disease; Parkinson's disease; psychosis; neurosis;

KW immunological deficiency; cancer therapy; stimulantion; modulator;

KW dietary supplement; cachexia; inhibition.

XX

OS Homo sapiens.  
 XX  
 PN WO9814473-A1.  
 XX  
 PD 09-APR-1998.  
 XX  
 PF 03-OCT-1997; 97WO-GB002721.  
 XX  
 PR 03-OCT-1996; 96PL-00316416.  
 XX  
 PA (HIRS-) HIRSZFELD INST IMMUNOLOGY & EXPERIMENTAL.  
 PA (GEOR-) GEORGIADIS BIOTECH LTD.  
 XX  
 PI Janusz M, Lisowski J, Dubowska-Inglot A;  
 XX  
 DR WPI; 1998-250967/22.  
 XX  
 PT Use of colostrinin, or derived nonapeptide for treating chronic diseases  
 PT of the central nervous system - and immune system, also as dietary  
 PT supplement and for inhibiting development of Alzheimer's disease.  
 XX  
 PS Claim 51; Page 27; 34pp; English.  
 XX  
 CC This is the amino acid sequence of the colostrinin derived nanopeptide  
 CC (NP). In the method of the invention colostrinin, and its NP are used to  
 CC treat chronic disorders of the central nervous system (CNS), particularly  
 CC neurological and mental disorders such as dementia (Alzheimer's disease);  
 CC motor neurone disease (e.g. Parkinson's disease); psychosis and neurosis  
 CC (including assisting withdrawal from addictive drugs) and the immune  
 CC system, particularly bacterial or viral infections or acquired  
 CC immunological deficiency (e.g. where caused by cancer therapy).  
 CC Colostrinin is a stimulant/modulator of the immune system and may also be  
 CC used as a dietary supplement, e.g. in babies who have not received  
 CC colostrum, young children or adults being given chemotherapy or suffering  
 CC from cachexia due to chronic disease, and to inhibit development of  
 CC Alzheimer's disease  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 2; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
 |||||  
 Db 1 VESYVPLFP 9

RESULT 2  
 AAB72529  
 ID AAB72529 standard; peptide; 9 AA.  
 XX  
 AC AAB72529;  
 XX  
 DT 09-MAY-2001 (first entry)  
 XX  
 DE Colostrinin peptide #30.

XX  
 KW Dermatological; oxidative stress regulator; colostrinin.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200112650-A2.  
 XX  
 PD 22-FEB-2001.  
 XX  
 PF 17-AUG-2000; 2000WO-US022665.  
 XX  
 PR 17-AUG-1999; 99US-0149310P.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX  
 PI Stanton GJ, Hughes TK, Boldogh I;  
 XX  
 DR WPI; 2001-218342/22.  
 XX  
 PT Modulating oxidative stress level in a cell, involves contacting the cell  
 PT with an oxidative stress regulator selected from colostrinin, its  
 PT constituent peptide, analog or their combinations.  
 XX  
 PS Claim 6; Page 26; 48pp; English.  
 XX  
 CC The present invention relates to a method for modulating the oxidative  
 CC stress level in a cell or a patient, comprising contacting the cell with,  
 CC or administering to the patient, an oxidative stress regulator selected  
 CC from colostrinin, or its constituent peptide (e.g. the present peptide),  
 CC to change the level of an oxidising species in the cell. The method can  
 CC be used to treat oxidative damage to skin, by decreasing or preventing an  
 CC increase in the level of damage to a biomolecule of the patient  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 4; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
 |||||  
 Db 1 VESYVPLFP 9

# RESULT 3

AAB72276

ID AAB72276 standard; peptide; 9 AA.

XX

AC AAB72276;

XX

DT 14-MAY-2001 (first entry)

XX

DE Colostrinin derived cytokine inducing peptide SEQ ID 31.

XX

KW Colostrinin; immune response; cytokine; blood cell proliferation;

KW central nervous system disorder; neurological disorder; mental disorder;

KW dementia; neurodegenerative disease; Alzheimer's disease; psychosis;

KW neurosis; infection.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200111937-A2.  
 XX  
 PD 22-FEB-2001.  
 XX  
 PF 17-AUG-2000; 2000WO-US022818.  
 XX  
 PR 17-AUG-1999; 99US-0149311P.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 PA (REGE-) REGEN THERAPEUTICS PLC.  
 XX  
 PI Stanton GJ, Hughes TK, Boldogh I, Georgiades J;  
 XX  
 DR WPI; 2001-202804/20.  
 XX  
 PT Inducing a cytokine and modulating an immune response, useful for  
 PT treating central nervous system diseases and bacterial and viral  
 PT infections, comprises administering colostrinin as an immunological  
 PT regulator.  
 XX  
 PS Claim 1; Page 34; 50pp; English.  
 XX  
 CC Sequences AAB72246 - AAB72275 represent peptides derived from clostrinin,  
 CC a proline rich polypeptide aggregate contained in colostrum. The peptides  
 CC have immune response modulatory activity, and are capable of inducing  
 CC cytokines. Colostrinin and its derived peptides are useful for inducing  
 CC cytokine production, for modulating an immunological response and for  
 CC inducing blood cell proliferation. The peptides are useful in the  
 CC treatment of disorders of the central nervous system, neurological  
 CC disorders, mental disorders, dementia, neurodegenerative diseases,  
 CC Alzheimer's disease, motor neurone disease, psychosis, neurosis, chronic  
 CC disorders of the immune system, bacterial and viral infections and  
 CC acquired immunological deficiencies  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 4; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
 |||||  
 Db 1 VESYVPLFP 9

RESULT 4  
 AAB72561  
 ID AAB72561 standard; peptide; 9 AA.  
 XX  
 AC AAB72561;  
 XX  
 DT 09-MAY-2001 (first entry)  
 XX

DE Colostrinin peptide #30.  
 XX  
 KW Neuroprotective; neural cell differentiation regulator; colostrinin;  
 KW colostrum.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200112651-A2.  
 XX  
 PD 22-FEB-2001.  
 XX  
 PF 17-AUG-2000; 2000WO-US022774.  
 XX  
 PR 17-AUG-1999; 99US-0149633P.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX  
 PI Boldogh I;  
 XX  
 DR WPI; 2001-226545/23.  
 XX  
 PT Use of colostrinin, its constituent peptide or analog as a neural cell  
 PT regulator, for promoting neural cell differentiation and treating damaged  
 PT neural cells in a patient.  
 XX  
 PS Claim 6; Page 22; 35pp; English.  
 XX  
 CC The present invention relates to a method for promoting neural cell  
 CC differentiation and treating damaged neural cells, using colostrinin and  
 CC colostrinin constituent peptides (e.g. the present peptide) as a neural  
 CC cell regulator. Colostrinin is a polypeptide complex found in colostrum  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 4; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
 |||||  
 Db 1 VESYVPLFP 9

# RESULT 5

AAO14607

ID AAO14607 standard; peptide; 9 AA.

XX

AC AAO14607;

XX

DT 27-MAY-2002 (first entry)

XX

DE Neural cell regulatory colostrinin peptide 30.

XX

KW Neural cell differentiation; neural cell regulator; colostrinin peptide;

KW neural cell formation; proline-rich polypeptide aggregate; colostrum;

KW neural cell treatment.

XX

OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 9  
 FT /note= "Optional C-terminal amide"  
 XX  
 PN WO200213851-A1.  
 XX  
 PD 21-FEB-2002.  
 XX  
 PF 17-AUG-2000; 2000WO-US022777.  
 XX  
 PR 17-AUG-2000; 2000WO-US022777.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX  
 PI Boldogh I, Stanton JG, Hughes TK;  
 XX  
 DR WPI; 2002-269152/31.  
 XX  
 PT Promoting cell differentiation in a patient involves use of blood cell  
 PT regulator selected from colostrinin, its constituent peptide and/or  
 PT analog.  
 XX  
 PS Claim 7; Page 22; 37pp; English.  
 XX  
 CC The invention comprises a method for promoting cell differentiation (e.g.  
 CC neural cell differentiation). The method involves contacting cells with a  
 CC neural cell regulator (i.e. a colostrinin peptide) in order to change the  
 CC cells in morphology to form neural cells. Colostrinin is a proline-rich  
 CC polypeptide aggregate that is present in colostrum. The method of the  
 CC invention is useful for promoting the differentiation of cells and for  
 CC treating damaged neural cells in a patient. The present amino acid  
 CC sequence represents a specifically claimed colostrinin peptide used in  
 CC the method of the invention  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 5; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
 |||||  
 Db 1 VESYVPLFP 9

RESULT 6  
 AAM51067  
 ID AAM51067 standard; peptide; 9 AA.  
 XX  
 AC AAM51067;  
 XX  
 DT 30-MAY-2002 (first entry)  
 XX  
 DE Colostrinin constituent peptide.  
 XX

KW Colostrinin; colostrum; beta-casein; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200213849-A1.  
 XX  
 PD 21-FEB-2002.  
 XX  
 PF 17-AUG-2000; 2000WO-US022775.  
 XX  
 PR 17-AUG-2000; 2000WO-US022775.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 PA (REGE-) REGEN THERAPEUTICS PLC.  
 XX  
 PI Stanton GJ, Hughes TK, Boldogh I, Georgiades J;  
 XX  
 DR WPI; 2002-269150/31.  
 XX  
 PT Modulation of blood cell proliferation in a patient involves use of blood  
 PT cell regulator selected from colostrinin, its constituent peptide and/or  
 PT analog.  
 XX  
 PS Example 1; Page 34; 54pp; English.  
 XX  
 CC The present sequence is that of a colostrinin constituent peptide that  
 CC has been classified as having a beta-casein homologue precursor. The  
 CC peptide has previously been identified as being useful in the treatment  
 CC of disorders of the central nervous system, neurological disorders,  
 CC mental disorders, dementia, neurodegenerative diseases, Alzheimer's  
 CC disease, motor neurone disease, psychosis, neurosis, chronic disorders of  
 CC the immune system, diseases with a bacterial or viral aetiology, and  
 CC acquired immunological deficiencies. The present invention provides  
 CC claimed colistrinin constituted peptides (see AAM51036-66) that are  
 CC useful as immunological regulators and as blood cell regulators. These  
 CC are used in claimed methods of the invention to modulate specific or  
 CC nonspecific immune responses in patients, and to modulate cellular  
 CC proliferation or differentiation of blood cells, such as leucocytes.  
 CC These methods specifically do not use the present colostrinin peptide  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 5; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
 |||||  
 Db 1 VESYVPLFP 9

RESULT 7  
 AAE20258  
 ID AAE20258 standard; peptide; 9 AA.  
 XX  
 AC AAE20258;  
 XX



DT 18-JUN-2002 (first entry)  
 XX  
 DE Colostrinin constituent peptide #30.  
 XX  
 KW Blood cell regulator; colostrinin; constituent peptide; oxidative stress;  
 KW therapy; oxidative damage; skin; aging; wound healing; cell replacement;  
 KW tissue; organ; cosmetic procedure; repair; regeneration; preservation;  
 KW transplantation; implantation; dermatological; vulnerary.  
 XX  
 OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 9  
 FT /note= "Optionally C-terminal amide"  
 XX  
 PN WO200213850-A1.  
 XX  
 PD 21-FEB-2002.  
 XX  
 PF 17-AUG-2000; 2000WO-US022776.  
 XX  
 PR 17-AUG-2000; 2000WO-US022776.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX  
 PI Stanton GJ, Hughes TK, Boldogh I;  
 XX  
 DR WPI; 2002-269151/31.  
 XX  
 PT Composition useful for the modulation of blood cell proliferation in a  
 PT patient comprises a blood cell regulator selected from colostrinin, its  
 PT constituent peptide and/or analog.  
 XX  
 PS Claim 6; Page 26; 51pp; English.  
 XX  
 CC The invention relates to a composition which comprises a blood cell  
 CC regulator selected from colostrinin, its constituent peptide and/or  
 CC analogue. The invention is used for modulating the oxidative stress level  
 CC in a cell e.g. mammalian or human cell present in a cell culture, tissue,  
 CC organ, or organism; or for treating oxidative damage to the skin of a  
 CC patient e.g. animal or human; to modulate oxidative stress during/ after  
 CC a premature birth or normal birth, preventing/delaying aging in a  
 CC patient, enhancing wound healing, and the reduction of side effects of  
 CC cosmetic procedures. The method changes the level of an oxidising species  
 CC in the cell, such as decreases or prevents increase in the level of  
 CC damage to a biomolecule of the patient selected from DNA, protein and/or  
 CC lipid, compared to the same conditions when the oxidative stress  
 CC regulator is not present. The modulation of oxidative stress results in  
 CC enhanced repair, regeneration, and replacement of cells, tissues and  
 CC organs (e.g. kidney, liver, pancreas, skin, and the other internal and  
 CC external organs), as well as enhanced preservation of such organs for  
 CC transplantation, implantation, or scientific research. The present  
 CC sequence is a colostrinin constituent peptide  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 5; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
|||||||  
Db 1 VESYVPLFP 9

RESULT 8

AAE07189

ID AAE07189 standard; peptide; 8 AA.

XX

AC AAE07189;

XX

DT 06-NOV-2001 (first entry)

XX

DE Colostrinin peptide 5.

XX

KW Colostrinin; nootropic; neuroprotective; immunomodulatory; antibacterial;

KW Parkinson's disease; Alzheimer's disease; mental disorder; food additive;

KW central nervous system disorder; neurodegenerative disorder; weight loss;

KW beta-amyloid plaque; psychosis; neurosis; cachexia; motor neuron disease;

KW acquired immunological deficiency; neurological disorder; dementia;

KW antiviral.

XX

OS Unidentified.

XX

PN WO200155199-A1.

XX

PD 02-AUG-2001.

XX

PF 26-JAN-2001; 2001WO-GB000329.

XX

PR 26-JAN-2000; 2000GB-00001825.

XX

PA (REGE-) REGEN THERAPEUTICS PLC.

XX

PI Georgiades JA;

XX

DR WPI; 2001-488775/53.

XX

PT Peptide useful as an inter alia in the treatment of e.g. disorders of the

PT immune system and the central nervous system comprises ten amino-terminal

PT amino acid sequence derived from peptides present in colostrinin.

XX

PS Claim 1; Page 15; 40pp; English.

XX

CC The invention relates to colostrinin peptide fragments which are useful,

CC inter alia, in the treatment of chronic disorders of the immune system

CC and the central nervous system. Colostrinin peptides are used as a

CC medicament in the treatment of neurological disorders e.g., dementia,

CC neurodegenerative disorders e.g., Alzheimer's disease, motor neuron

CC disease e.g., Parkinson's disease, mental disorders e.g. psychosis and

CC neurosis, in acquired immunological deficiencies, chronic bacterial and

CC viral infections and diseases characterised by the presence of beta-

CC amyloid plaques and as a dietary supplement for babies, small children,

CC adults and senile persons, who have been subjected to chemotherapy or

CC have suffered from cachexia or weight loss due to the chronic disease.  
CC Colostrinin peptides are also used as food additives and as an auxillary  
CC withdrawal treatment for drug addicts, after a period of detoxification  
CC and in persons dependent on stimulants. Colostrinin peptides are used to  
CC prepare antibodies and to treat emotional disturbances, e.g. emotional  
CC disturbances of psychiatric patients in a state of depression. These  
CC colostrinin peptides improves the development of immune system in a new  
CC born child and to correct the immunological deficiencies in a child. The  
CC present sequence is colostrinin peptide 5 related to the invention  
XX  
SQ Sequence 8 AA;

Query Match 91.7%; Score 44; DB 4; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9  
| | | | | | | |  
Db 1 ESYVPLFP 8

RESULT 9

AAE07199

ID AAE07199 standard; peptide; 9 AA.

XX

AC AAE07199;

XX

DT 06-NOV-2001 (first entry)

XX

DE Modified colostrinin cyclic peptide #5.

XX

KW Colostrinin; nootropic; neuroprotective; immunomodulatory; antibacterial;  
KW Parkinson's disease; Alzheimer's disease; mental disorder; food additive;  
KW central nervous system disorder; neurodegenerative disorder; weight loss;  
KW beta-amyloid plaque; psychosis; neurosis; cachexia; motor neuron disease;  
KW acquired immunological deficiency; neurological disorder; dementia;  
KW antiviral; cyclic.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Modified-site 1

FT /note= "N-terminal acetyl; this residue forms a cyclic  
FT linkage with Pro found at the C-terminal end"

XX

PN WO200155199-A1.

XX

PD 02-AUG-2001.

XX

PF 26-JAN-2001; 2001WO-GB000329.

XX

PR 26-JAN-2000; 2000GB-00001825.

XX

PA (REGE-) REGEN THERAPEUTICS PLC.

XX

PI Georgiades JA;

XX

DR WPI; 2001-488775/53.

XX

PT Peptide useful as an inter alia in the treatment of e.g. disorders of the  
PT immune system and the central nervous system comprises ten amino-terminal  
PT amino acid sequence derived from peptides present in colostrinin.

XX

PS Example 2; Page 9; 40pp; English.

XX

CC The invention relates to colostrinin peptide fragments which are useful,  
CC inter alia, in the treatment of chronic disorders of the immune system  
CC and the central nervous system. Colostrinin peptides are used as a  
CC medicament in the treatment of neurological disorders e.g., dementia,  
CC neurodegenerative disorders e.g., Alzheimer's disease, motor neuron  
CC disease e.g., Parkinson's disease, mental disorders e.g. psychosis and  
CC neurosis, in acquired immunological deficiencies, chronic bacterial and  
CC viral infections and diseases characterised by the presence of beta-  
CC amyloid plaques and as a dietary supplement for babies, small children,  
CC adults and senile persons, who have been subjected to chemotherapy or  
CC have suffered from cachexia or weight loss due to the chronic disease.  
CC Colostrinin peptides are also used as food additives and as an auxillary  
CC withdrawal treatment for drug addicts, after a period of detoxification  
CC and in persons dependent on stimulants. Colostrinin peptides are used to  
CC prepare antibodies and to treat emotional disturbances, e.g. emotional  
CC disturbances of psychiatric patients in a state of depression. These  
CC colostrinin peptides improves the development of immune system in a new  
CC born child and to correct the immunological deficiencies in a child. The  
CC present sequence is modified colostrinin cyclic peptide #5 related to the  
CC invention

XX

SQ Sequence 9 AA;

Query Match 91.7%; Score 44; DB 4; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9

|||||||

Db 2 ESYVPLFP 9

RESULT 10

AAB18662

ID AAB18662 standard; protein; 249 AA.

XX

AC AAB18662;

XX

DT 22-JAN-2001 (first entry)

XX

DE A human regulator of intracellular phosphorylation.

XX

KW Human; intracellular phosphorylation regulator; HRIP; stroke; myeloma;

KW neurological disorder; Parkinson's disease; demyelinating disease;

KW meningitis; developmental disorder; neuromuscular disorder; cancer;

KW myasthenia gravis; cell proliferative disorder; actinic keratosis;

KW arteriosclerosis; atherosclerosis; leukaemia; melanoma; bronchitis;

KW autoimmune disorder; inflammatory disorder; Addison's disease;

KW acquired immunodeficiency disease; allergy; diabetes mellitus;

KW    rheumatoid arthritis; microbial infection; trauma.  
XX  
OS    Homo sapiens.  
XX  
FH    Key                    Location/Qualifiers  
FT    Modified-site        3  
FT                            /note= "potential phosphorylation site"  
FT    Modified-site        4  
FT                            /note= "potential phosphorylation site"  
FT    Modified-site        38  
FT                            /note= "potential phosphorylation site"  
FT    Modified-site        64  
FT                            /note= "potential phosphorylation site"  
FT    Modified-site        75  
FT                            /note= "potential phosphorylation site"  
FT    Active-site          88. .100  
FT                            /note= "tyrosine specific protein phosphatase active  
FT                            site"  
FT    Modified-site        107  
FT                            /note= "potential phosphorylation site"  
FT    Modified-site        119  
FT                            /note= "potential phosphorylation site"  
FT    Modified-site        137  
FT                            /note= "potential phosphorylation site"  
FT    Modified-site        150  
FT                            /note= "potential phosphorylation site"  
FT    Modified-site        196  
FT                            /note= "potential phosphorylation site"  
FT    Modified-site        204  
FT                            /note= "potential glycosylation site"  
XX  
PN    WO200055332-A2.  
XX  
PD    21-SEP-2000.  
XX  
PF    17-MAR-2000; 2000WO-US007277.  
XX  
PR    18-MAR-1999;        99US-0125593P.  
PR    20-MAY-1999;        99US-0135049P.  
PR    09-JUL-1999;        99US-0143188P.  
XX  
PA    (INCY-) INCYTE PHARM INC.  
XX  
PI    Bandman O,   Tang YT,   Yue H,   Hillman JL,   Baughn MR,   Azimzai Y;  
PI    Lu DAM,    Au-Young J;  
XX  
DR    WPI; 2000-602121/57.  
DR    N-PSDB; AAA75679.  
XX  
PT    Novel human intracellular phosphorylation regulator polypeptides and  
PT    polynucleotides for diagnosis, prevention and treatment of neurological,  
PT    cell proliferative and autoimmune/inflammatory disorders.  
XX  
PS    Claim 1; Page 80; 96pp; English.  
XX  
CC    The present sequence represents a human regulator of intracellular  
CC    phosphorylation (HRIP). HRIP is useful for screening agonists and

CC antagonists of HRIP polypeptide. HRIP and its agonist or antagonist are  
CC useful for treating a disease or condition associated with decreased or  
CC increased expression of functional HRIP. Diseases treated or diagnosed  
CC include neurological disorders such as stroke, Parkinson's disease,  
CC demyelinating diseases, bacterial and viral meningitis and other  
CC developmental disorders of the central nervous system, neuromuscular  
CC disorders, myasthenia gravis, cell proliferative disorders such as  
CC actinic keratosis, arteriosclerosis, atherosclerosis and cancer including  
CC leukaemia, melanoma, myeloma and cancer of the adrenal gland, bladder,  
CC bone, bone marrow, liver, lung, muscle, ovary, autoimmune/ inflammatory  
CC disorder such as Addison's disease, acquired immunodeficiency diseases,  
CC allergies, bronchitis, diabetes mellitus, rheumatoid arthritis, microbial  
CC infection and trauma

XX

SQ Sequence 249 AA;

Query Match 79.2%; Score 38; DB 3; Length 249;

Best Local Similarity 85.7%; Pred. No. 57;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
||:||||  
Db 219 SYIPLFP 225

#### RESULT 11

AAY78810

ID AAY78810 standard; protein; 249 AA.

XX

AC AAY78810;

XX

DT 09-MAY-2000 (first entry)

XX

DE Hydrophobic domain containing protein clone HP10551 protein sequence.

XX

KW Hydrophobic domain; clone HP10551; nutritional supplement; SCID; HIV;  
KW cell proliferation; immune stimulant; immune deficiency; tumour; pain;  
KW rheumatoid arthritis; insulin dependent diabetes mellitus; fertility;  
KW myasthenia gravis; haematopoiesis regulator; tissue growth; depression;  
KW anti-inflammatory; infection; bodily characteristic.

XX

OS Homo sapiens.

XX

PN WO200000506-A2.

XX

PD 06-JAN-2000.

XX

PF 18-JUN-1999; 99WO-JP003242.

XX

PR 26-JUN-1998; 98JP-00180008.

XX

PA (SAGA ) SAGAMI CHEM RES CENT.

PA (PROT-) PROTEGENE INC.

XX

PI Kato S, Kimura T;

XX

DR WPI; 2000-160665/14.

DR N-PSDB; AAZ90047, AAZ90057.

XX

PT Novel human proteins having hydrophobic domains used for research and  
PT diagnostic purposes.

XX

PS Claim 1; Page 82-83; 117pp; English.

XX

CC This sequence represents the hydrophobic domain containing protein, clone  
CC HP10551 protein sequence. The sequence is isolated from a human stomach  
CC cancer cell line. The invention relates to human proteins with  
CC hydrophobic domains, the DNA and the cDNA encoding them. The  
CC polynucleotides and proteins are predicted to have biological activities  
CC which make them suitable for treating, preventing or ameliorating medical  
CC conditions in humans and animals. Suggested activities include  
CC nutritional activity (nutritional source or supplement); cytokine and  
CC cell proliferation/differentiation activity; immune stimulating (e.g. as  
CC vaccines) or suppressing activity (e.g. to treat various immune  
CC deficiencies such as SCIDS or HIV, connective tissue disease, systemic  
CC lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary  
CC inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin  
CC dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease  
CC and autoimmune inflammatory eye disease, as well as asthma, allergies and  
CC organ transplantation); haematopoiesis regulating activity (e.g. in  
CC treatment of myeloid or lymphoid cell deficiencies); tissue growth  
CC activity (e.g. wound healing and tissue repair, ulcers, burns,  
CC periodontal disease); activin/inhibin activity; chemotactic/chemokinetic  
CC activity; haemostatic and thrombolytic activity (e.g. treating  
CC haemophilias); receptor/ligand activity; anti-inflammatory activity; and  
CC tumour inhibition activity. The polynucleotides are also stated to be  
CC useful for gene therapy. Other activities include inhibiting infections  
CC caused by bacteria, fungi, viruses and other parasites (e.g. Hepatitis,  
CC malaria); effecting bodily characteristics such as, e.g. weight, colour,  
CC skin, effecting biorhythms or circadian cycles; enhancing fertility;  
CC treatment of depression; treatment of pain; hormonal or endocrine  
CC activity. The polynucleotides may also be used for recombinant expression  
CC of the protein

XX

SQ Sequence 249 AA;

Query Match 79.2%; Score 38; DB 3; Length 249;

Best Local Similarity 85.7%; Pred. No. 57;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9

||:||||

Db 219 SYIPLFP 225

RESULT 12

AAM39162

ID AAM39162 standard; protein; 249 AA.

XX

AC AAM39162;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polypeptide SEQ ID NO 2307.

XX  
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000WO-US034263.  
 XX  
 PR 23-DEC-1999; 99US-00471275.  
 PR 21-JAN-2000; 2000US-00488725.  
 PR 25-APR-2000; 2000US-00552317.  
 PR 20-JUN-2000; 2000US-00598042.  
 PR 19-JUL-2000; 2000US-00620312.  
 PR 03-AUG-2000; 2000US-00653450.  
 PR 14-SEP-2000; 2000US-00662191.  
 PR 19-OCT-2000; 2000US-00693036.  
 PR 29-NOV-2000; 2000US-00727344.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
 PI Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI; 2001-442253/47.  
 DR N-PSDB; AAI58318.  
 XX  
 PT Novel nucleic acids and polypeptides, useful for treating disorders such  
 PT as central nervous system injuries.  
 XX  
 PS Example 4; SEQ ID NO 2307; 10078pp; English.  
 XX  
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
 CC encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders. Note: The sequence data for this patent did not form  
 CC part of the printed specification  
 XX  
 SQ Sequence 249 AA;



Query Match 79.2%; Score 38; DB 4; Length 249;  
Best Local Similarity 85.7%; Pred. No. 57;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
||:||||  
Db 219 SYIPLFP 225

RESULT 13

AAB42118

ID AAB42118 standard; protein; 288 AA.

XX

AC AAB42118;

XX

DT 08-FEB-2001 (first entry)

XX

DE Human ORFX ORF1882 polypeptide sequence SEQ ID NO:3764.

XX

KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
KW vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
KW cholesterol ester storage; systemic lupus erythematosus; infection;  
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
KW thrombosis; contraceptive.

XX

OS Homo sapiens.

XX

PN WO200058473-A2.

XX

PD 05-OCT-2000.

XX

PF 31-MAR-2000; 2000WO-US008621.

XX

PR 31-MAR-1999; 99US-0127607P.

PR 02-APR-1999; 99US-0127636P.

PR 05-APR-1999; 99US-0127728P.

PR 30-MAR-2000; 2000US-00540763.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Shimkets RA, Leach M;

XX

DR WPI; 2000-602362/57.

DR N-PSDB; AAC76327.

XX

PT Novel nucleic acids and peptides derived from open reading frame X,  
PT useful for treating e.g. cancers, proliferative disorders,

PT neurodegenerative disorders and cardiovascular disease.  
 XX  
 PS Claim 11; Page 2904-2905; 5507pp; English.  
 XX  
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human OREFX open reading frames 1 to 3161. The OREFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;  
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;  
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;  
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;  
 CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The  
 CC sequences can be used for determining the presence of or predisposition  
 CC to, or preventing or treating pathological conditions associated with an  
 CC OREFX-associated disorder. The nucleic acids can be used to express OREFX  
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
 CC used to treat cancers, proliferative disorders, neurodegenerative  
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
 XX  
 SQ Sequence 288 AA;

Query Match 79.2%; Score 38; DB 3; Length 288;  
 Best Local Similarity 85.7%; Pred. No. 66;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
 ||:||||  
 Db 258 SYIPLFP 264

#### RESULT 14

AAG81301

ID AAG81301 standard; protein; 288 AA.

XX

AC AAG81301;

XX

DT 10-SEP-2001 (first entry)

XX

DE Human AFP protein sequence SEQ ID NO:120.

XX

KW Human; secreted protein; secretion; bacterial cell; fungal cell;  
 KW eukaryotic cell; fusion protein; maltose binding protein;  
 KW immunoglobulin constant region; polyhistidine tag.

XX

OS Homo sapiens.

XX

PN WO200129221-A2.

XX

PD 26-APR-2001.

XX

PF 20-OCT-2000; 2000WO-US029052.

XX  
 PR 20-OCT-1999; 99US-0160712P.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Conklin DC, Yee DP;  
 XX  
 DR WPI; 2001-300340/31.  
 DR N-PSDB; AAH52152.  
 XX  
 PT Isolated polypeptide for directing secretion of proteins of interest from  
 PT a host cell including, e.g. bacteria, includes contiguous amino acid  
 PT residues of polypeptide with specified amino acids.  
 XX  
 PS Claim 1; Page 239-240; 617pp; English.  
 XX  
 CC AAH52093 to AAH52303 encode the human secreted proteins given in AAG81242  
 CC to AAG81453. The secreted proteins can be used for directing the  
 CC secretion of proteins of interest from a host cell including bacteria,  
 CC fungal cells, and cultured higher eukaryotic cells. The present invention  
 CC also describes fusion proteins, where a secreted protein of the invention  
 CC is operably linked via a peptide bond or peptide linker to a second  
 CC protein selected from the group consisting of maltose binding protein, an  
 CC immunoglobulin constant region, a polyhistidine tag and a peptide given  
 CC in AAG81453  
 XX  
 SQ Sequence 288 AA;

Query Match 79.2%; Score 38; DB 4; Length 288;  
 Best Local Similarity 85.7%; Pred. No. 66;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
 ||:||||  
 Db 258 SYIPLFP 264

# RESULT 15

AAM40948

ID AAM40948 standard; protein; 326 AA.

XX

AC AAM40948;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polypeptide SEQ ID NO 5879.

XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.

XX

OS Homo sapiens.

XX

PN WO200153312-A1.

XX  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000WO-US034263.  
 XX  
 PR 23-DEC-1999; 99US-00471275.  
 PR 21-JAN-2000; 2000US-00488725.  
 PR 25-APR-2000; 2000US-00552317.  
 PR 20-JUN-2000; 2000US-00598042.  
 PR 19-JUL-2000; 2000US-00620312.  
 PR 03-AUG-2000; 2000US-00653450.  
 PR 14-SEP-2000; 2000US-00662191.  
 PR 19-OCT-2000; 2000US-00693036.  
 PR 29-NOV-2000; 2000US-00727344.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
 PI Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI; 2001-442253/47.  
 DR N-PSDB; AAI60104.  
 XX  
 PT Novel nucleic acids and polypeptides, useful for treating disorders such  
 PT as central nervous system injuries.  
 XX  
 PS Example 2; SEQ ID NO 5879; 10078pp; English.  
 XX  
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
 CC encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders. Note: The sequence data for this patent did not form  
 CC part of the printed specification  
 XX  
 SQ Sequence 326 AA;

Query Match 79.2%; Score 38; DB 4; Length 326;  
 Best Local Similarity 85.7%; Pred. No. 75;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
 ||:||||  
 Db 256 SYIPLFP 262

Search completed: August 24, 2004, 15:42:47

Job time : 42.6716 secs

OM protein - protein search, using sw model

Run on: August 24, 2004, 15:33:13 ; Search time 9.87313 Seconds  
(without alignments)  
47.060 Million cell updates/sec

Title: US-09-641-801-31  
Perfect score: 48  
Sequence: 1 VESYVPLFP 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	48	100.0	9	4	US-09-641-803-31	Sequence 31, Appl
2	36	75.0	301	4	US-09-540-236-2519	Sequence 2519, Ap
3	35	72.9	97	4	US-09-198-452A-71	Sequence 71, Appl
4	35	72.9	470	4	US-09-134-000C-4816	Sequence 4816, Ap
5	35	72.9	614	4	US-09-328-352-6661	Sequence 6661, Ap
6	35	72.9	1041	3	US-08-898-978-2	Sequence 2, Appli
7	35	72.9	1041	4	US-09-372-858-2	Sequence 2, Appli
8	34	70.8	394	4	US-09-252-991A-28427	Sequence 28427, A
9	34	70.8	989	3	US-09-110-517-4	Sequence 4, Appli
10	33	68.8	68	4	US-09-621-976-4917	Sequence 4917, Ap
11	33	68.8	289	3	US-09-077-675A-5	Sequence 5, Appli

12	33	68.8	289	4	US-09-077-674-5	Sequence 5, Appli
13	33	68.8	353	3	US-09-077-675A-3	Sequence 3, Appli
14	33	68.8	353	4	US-09-077-674-3	Sequence 3, Appli
15	33	68.8	366	4	US-09-762-661A-7	Sequence 7, Appli
16	33	68.8	366	4	US-09-743-475-6	Sequence 6, Appli
17	32	66.7	208	4	US-09-252-991A-31553	Sequence 31553, A
18	32	66.7	250	4	US-09-702-705-1677	Sequence 1677, Ap
19	32	66.7	250	4	US-09-736-457-1677	Sequence 1677, Ap
20	32	66.7	250	4	US-09-671-325-1677	Sequence 1677, Ap
21	32	66.7	359	4	US-09-328-352-7543	Sequence 7543, Ap
22	32	66.7	384	2	US-08-872-302-4	Sequence 4, Appli
23	32	66.7	385	4	US-09-198-452A-1092	Sequence 1092, Ap
24	32	66.7	391	3	US-09-347-798-8	Sequence 8, Appli
25	32	66.7	548	3	US-08-903-139B-28	Sequence 28, Appl
26	32	66.7	700	4	US-09-266-225D-6	Sequence 6, Appli
27	32	66.7	754	2	US-08-941-262-1	Sequence 1, Appli
28	31	64.6	78	4	US-09-134-001C-3761	Sequence 3761, Ap
29	31	64.6	82	4	US-09-673-395A-347	Sequence 347, App
30	31	64.6	173	4	US-09-543-681A-7808	Sequence 7808, Ap
31	31	64.6	195	4	US-09-198-452A-126	Sequence 126, App
32	31	64.6	409	4	US-09-540-236-2952	Sequence 2952, Ap
33	31	64.6	439	4	US-09-489-039A-13830	Sequence 13830, A
34	31	64.6	839	4	US-09-751-687-15	Sequence 15, Appl
35	31	64.6	1036	4	US-09-489-039A-10266	Sequence 10266, A
36	31	64.6	1159	4	US-09-328-352-7624	Sequence 7624, Ap
37	31	64.6	1185	4	US-09-252-991A-18328	Sequence 18328, A
38	30.5	63.5	35	1	US-08-487-890A-65	Sequence 65, Appl
39	30.5	63.5	35	2	US-08-478-435-65	Sequence 65, Appl
40	30.5	63.5	35	2	US-08-337-483-65	Sequence 65, Appl
41	30.5	63.5	35	2	US-08-478-373-65	Sequence 65, Appl
42	30.5	63.5	35	3	US-08-474-671-65	Sequence 65, Appl
43	30.5	63.5	35	3	US-08-483-577A-65	Sequence 65, Appl
44	30.5	63.5	35	3	US-08-897-438-65	Sequence 65, Appl
45	30.5	63.5	35	4	US-08-637-654-65	Sequence 65, Appl

# ALIGNMENTS

## RESULT 1

US-09-641-803-31

; Sequence 31, Application US/09641803

; Patent No. 6500798

; GENERAL INFORMATION:

; APPLICANT: STANTON, G. John

; APPLICANT: HUGHES, Thomas K.

; APPLICANT: BOLDOGH, Istvan

; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND

; TITLE OF INVENTION: ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS

; FILE REFERENCE: 265.00220101

; CURRENT APPLICATION NUMBER: US/09/641,803

; CURRENT FILING DATE: 2000-08-17

; PRIOR APPLICATION NUMBER: 60/149,310

; PRIOR FILING DATE: 1999-08-17

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 31

; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: peptide  
US-09-641-803-31

Query Match 100.0%; Score 48; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
| | | | | | | |  
Db 1 VESYVPLFP 9

RESULT 2

US-09-540-236-2519  
; Sequence 2519, Application US/09540236  
; Patent No. 6673910  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2005-001  
; CURRENT APPLICATION NUMBER: US/09/540,236  
; CURRENT FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 3840  
; SEQ ID NO 2519  
; LENGTH: 301  
; TYPE: PRT  
; ORGANISM: M.catarrhalis  
US-09-540-236-2519

Query Match 75.0%; Score 36; DB 4; Length 301;  
Best Local Similarity 75.0%; Pred. No. 50;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9  
| : | | | : |  
Db 274 EAYVPLWP 281

RESULT 3

US-09-198-452A-71  
; Sequence 71, Application US/09198452A  
; Patent No. 6559294  
; GENERAL INFORMATION:  
; APPLICANT: Griffais, R.  
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
; TITLE OF INVENTION: and treatment of infection  
; FILE REFERENCE: 9710-003-999



; CURRENT APPLICATION NUMBER: US/09/198,452A  
; CURRENT FILING DATE: 1998-11-24  
; NUMBER OF SEQ ID NOS: 6849  
; SEQ ID NO 71  
; LENGTH: 97  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-09-198-452A-71

Query Match 72.9%; Score 35; DB 4; Length 97;  
Best Local Similarity 75.0%; Pred. No. 23;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9  
|:|||||  
Db 1 ETYVPLLP 8

RESULT 4

US-09-134-000C-4816  
; Sequence 4816, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4816  
; LENGTH: 470  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-4816

Query Match 72.9%; Score 35; DB 4; Length 470;  
Best Local Similarity 71.4%; Pred. No. 1.2e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
||:|:|  
Db 376 SYIPVFP 382

RESULT 5

US-09-328-352-6661  
; Sequence 6661, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 6661  
; LENGTH: 614  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-6661

Query Match 72.9%; Score 35; DB 4; Length 614;  
Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9  
|||||: |  
Db 373 ESYVPVLP 380

RESULT 6

US-08-898-978-2  
; Sequence 2, Application US/08898978  
; Patent No. 6001602  
; GENERAL INFORMATION:  
; APPLICANT: Reichard, Raymond W.  
; APPLICANT: Brown, James R.  
; APPLICANT: Lawlor, Elizabeth J.  
; TITLE OF INVENTION: No. 6001602el ileS  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert Price & Rhoads  
; STREET: 997 Lenox Drive, Building 3, Suite 210  
; CITY: Lawrenceville  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 08543  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/898,978  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bloom, Allen  
; REGISTRATION NUMBER: 29,135  
; REFERENCE/DOCKET NUMBER: GM10051  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 609-520-3214  
; TELEFAX: 609-520-3259  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1041 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-898-978-2

Query Match 72.9%; Score 35; DB 3; Length 1041;  
Best Local Similarity 75.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9  
: || ||||  
Db 289 QSYEPLFP 296

RESULT 7

US-09-372-858-2

; Sequence 2, Application US/09372858  
; Patent No. 6361970  
; GENERAL INFORMATION:  
; APPLICANT: Reichard, Raymond  
; APPLICANT: Brown, James  
; APPLICANT: Lawlor, Elizabeth  
; TITLE OF INVENTION: NOVEL ileS  
; FILE REFERENCE: GM10051-D1  
; CURRENT APPLICATION NUMBER: US/09/372,858  
; CURRENT FILING DATE: 1999-12-08  
; PRIOR APPLICATION NUMBER: 08/898,978  
; PRIOR FILING DATE: 1997-07-23  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1041  
; TYPE: PRT  
; ORGANISM: Chlamydia trachomatis  
US-09-372-858-2

Query Match 72.9%; Score 35; DB 4; Length 1041;  
Best Local Similarity 75.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9  
: || ||||  
Db 289 QSYEPLFP 296

RESULT 8

US-09-252-991A-28427

; Sequence 28427, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 28427  
; LENGTH: 394  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-28427

Query Match 70.8%; Score 34; DB 4; Length 394;  
Best Local Similarity 62.5%; Pred. No. 1.5e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9  
| |::|  
Db 183 EIYIPLYP 190

RESULT 9

US-09-110-517-4  
; Sequence 4, Application US/09110517A  
; Patent No. 6248520  
; GENERAL INFORMATION:  
; APPLICANT: Roeder, Robert G  
; APPLICANT: Fondell, Joseph D  
; APPLICANT: Yuan, Chao X  
; APPLICANT: Ito, Mitsuhiro  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING NUCLEAR HORMONE  
; TITLE OF INVENTION: RECEPTOR COACTIVATORS AND USES THEREOF  
; FILE REFERENCE: 600-1-224  
; CURRENT APPLICATION NUMBER: US/09/110,517A  
; CURRENT FILING DATE: 1998-07-06  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 989  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-110-517-4

Query Match 70.8%; Score 34; DB 3; Length 989;  
Best Local Similarity 55.6%; Pred. No. 4e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
:| |::|  
Db 840 IEDYISLFP 848

RESULT 10

US-09-621-976-4917  
; Sequence 4917, Application US/09621976  
; Patent No. 6639063

; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 4917  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-621-976-4917

Query Match 68.8%; Score 33; DB 4; Length 68;  
Best Local Similarity 66.7%; Pred. No. 36;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
|| || ||  
Db 38 VEQYCPWFP 46

RESULT 11

US-09-077-675A-5

; Sequence 5, Application US/09077675A  
; Patent No. 6242199  
; GENERAL INFORMATION:  
; APPLICANT: Pai, Lee-Yuh  
; APPLICANT: Feighner, Scott C.  
; APPLICANT: Howard, Andrew D.  
; APPLICANT: Pong, Sheng-Shung  
; APPLICANT: Van Der Ploeg, Leonardus H.T.  
; TITLE OF INVENTION: RECEPTOR ASSAY  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merck & Co., Inc.  
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
; CITY: Rahway  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07065-0900  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/077,675A  
; FILING DATE: 3-JUN-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:

```

; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-675A-5

```

```

Query Match          68.8%; Score 33; DB 3; Length 289;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches      6; Conservative      1; Mismatches      2; Indels      0; Gaps      0;

```

```

Qy      1 VESYVPLFP 9
        || :|||
Db      31 VEELLPLFP 39

```

RESULT 12

US-09-077-674-5

```

; Sequence 5, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

```

```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-674-5

```

```

Query Match          68.8%; Score 33; DB 4; Length 289;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches      6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1 VESYVPLFP 9
        || :|||
Db      31 VEELLPLFP 39

```

RESULT 13

US-09-077-675A-3

```

; Sequence 3, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

```

```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-675A-3

```

```

Query Match          68.8%; Score 33; DB 3; Length 353;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches      6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1 VESYVPLFP 9
        || :||||
Db      18 VEELLPLFP 26

```

RESULT 14

US-09-077-674-3

```

; Sequence 3, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION: 536

```



```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-674-3

```

```

Query Match          68.8%; Score 33; DB 4; Length 353;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches      6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1 VESYVPLFP 9
        || :|||
Db      18 VEELLPLFP 26

```

RESULT 15

US-09-762-661A-7

```

; Sequence 7, Application US/09762661A
; Patent No. 6645726
; GENERAL INFORMATION:
; APPLICANT: Howard, Andrew D.
; APPLICANT: Palyha, Oksana C.
; APPLICANT: Smith, Roy G.
; APPLICANT: Tan, Carina P.
; TITLE OF INVENTION: CANINE GROWTH HORMONE SECRETAGOGUE
; TITLE OF INVENTION: RECEPTOR
; FILE REFERENCE: 20207P
; CURRENT APPLICATION NUMBER: US/09/762,661A
; CURRENT FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US99/17915
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 60/095,960
; PRIOR FILING DATE: 1998-08-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-762-661A-7

```

```

Query Match          68.8%; Score 33; DB 4; Length 366;
Best Local Similarity 66.7%; Pred. No. 2.1e+02;

```

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 VESYVPLFP 9  
|| :|||  
Db 31 VEELLPLFP 39

Search completed: August 24, 2004, 15:55:21  
Job time : 10.8731 secs

OM protein - protein search, using sw model

Run on: August 24, 2004, 15:26:28 ; Search time 8.73134 Seconds  
 (without alignments)  
 99.151 Million cell updates/sec

Title: US-09-641-801-31  
 Perfect score: 48  
 Sequence: 1 VESYVPLFP 9

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : PIR\_78:\*  
 1: pir1:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Query				ID	Description
		Match	Length	DB			
1	37	77.1	2136	2		A05037	hypothetical prote
2	36	75.0	417	2		D90304	amino acid transpo
3	35	72.9	301	2		S62459	hypothetical prote
4	35	72.9	337	2		T37731	probable dihydroor
5	35	72.9	466	2		C86837	amino acid antipor
6	35	72.9	471	2		H70559	hypothetical prote
7	35	72.9	1036	2		F71565	probable isoleucin
8	35	72.9	1365	2		T18419	hypothetical prote
9	34	70.8	181	2		D83688	hypothetical prote
10	34	70.8	202	2		AF2088	hypothetical prote
11	34	70.8	204	2		T10766	patatin-like latex
12	34	70.8	385	2		A40359	aliphatic amidase
13	34	70.8	385	2		C83226	aliphatic amidase

14	34	70.8	395	2	S66738	hypothetical prote
15	34	70.8	768	2	T38188	probable DNA repai
16	34	70.8	780	2	S62418	hypothetical prote
17	33	68.8	258	2	AE1819	hypothetical prote
18	33	68.8	320	2	B72656	hypothetical prote
19	33	68.8	351	2	T40854	hypothetical prote
20	33	68.8	392	2	F69544	conserved hypothet
21	33	68.8	439	2	T52304	phytochrome kinase
22	33	68.8	439	2	E84442	hypothetical prote
23	33	68.8	442	2	H86276	F14L17.4 protein -
24	33	68.8	457	2	G82925	hypothetical prote
25	33	68.8	516	2	AG0560	hypothetical prote
26	33	68.8	612	2	T45641	beta-D-glucan exoh
27	33	68.8	641	2	H69651	lichenan operon tr
28	33	68.8	713	2	JC2534	RVLG protein - rat
29	33	68.8	743	2	T40521	hypothetical prote
30	33	68.8	882	2	AE1416	sugar hydrolase ho
31	33	68.8	988	2	E96621	hypothetical prote
32	32	66.7	119	2	E97957	hypothetical prote
33	32	66.7	190	2	B83321	hypothetical prote
34	32	66.7	213	2	T46069	hypothetical prote
35	32	66.7	221	2	T09831	alcohol dehydrogen
36	32	66.7	233	2	B81666	lipoate-protein li
37	32	66.7	237	2	T40603	hypothetical prote
38	32	66.7	286	2	A48399	probable oxidoredu
39	32	66.7	286	2	A85739	probable dehydroge
40	32	66.7	286	2	H90879	probable dehydroge
41	32	66.7	288	2	AE3025	oxidoreductase Atu
42	32	66.7	306	2	A83369	probable transcrip
43	32	66.7	324	2	T11764	NADH2 dehydrogenas
44	32	66.7	334	2	D98259	hypothetical oxido
45	32	66.7	348	2	S44628	f22b7.1 protein -

# ALIGNMENTS

## RESULT 1

A05037

hypothetical protein 2136 - liverwort (*Marchantia polymorpha*) chloroplast

C;Species: chloroplast *Marchantia polymorpha*

C;Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 21-Jul-2000

C;Accession: S01591; A05037

R;Umesono, K.; Inokuchi, H.; Shiki, Y.; Takeuchi, M.; Chang, Z.; Fukuzawa, H.; Kohchi, T.; Shirai, H.; Ohyama, K.; Ozeki, H.

J. Mol. Biol. 203, 299-331, 1988

A;Title: Structure and organization of *Marchantia polymorpha* chloroplast genome.

II. Gene organization of the large single copy region from rps'12 to atpB.

A;Reference number: S01567; MUID:89068686; PMID:2974085

A;Accession: S01591

A;Molecule type: DNA

A;Residues: 1-2136 <UME>

A;Cross-references: EMBL:X04465; NID:g11640; PIDN:CAA28078.1; PID:g11665

R;Ohyama, K.; Fukuzawa, H.; Kohchi, T.; Shirai, H.; Sano, T.; Sano, S.; Umesono, K.; Shiki, Y.; Takeuchi, M.; Chang, Z.; Aota, S.; Inokuchi, H.; Ozeki, H.

Nature 322, 572-574, 1986

A;Title: Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA.

A;Reference number: A38014

A;Contents: annotation; gene organization, sites, features

C;Genetics:

A;Genome: chloroplast

C;Keywords: chloroplast

Query Match 77.1%; Score 37; DB 2; Length 2136;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ESYVPLF 8  
| | | | | | | |  
Db 1420 ESYVPLF 1426

#### RESULT 2

D90304

amino acid transporter SS01463 [imported] - *Sulfolobus solfataricus*

C;Species: *Sulfolobus solfataricus*

C;Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001

C;Accession: D90304

R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-Weiher, C.C.Y.; Clausen, I.G.; Curtis, B.A.; De Moors, A.; Erauso, G.; Fletcher, C.; Gordon, P.M.K.; Heikamp-de Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.; Schenk, M.E.; Theriault, C.; Tolstrup, N.; Charlebois, R.L.; Doolittle, W.F.; Duguet, M.; Gaasterland, T.; Garrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A;Description: *Sulfolobus solfataricus* complete genome.

A;Reference number: A99139

A;Accession: D90304

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-417 <KUR>

A;Cross-references: GB:AE006641; NID:g13814691; PIDN:AAK41691.1; GSPDB:GN00155

C;Genetics:

A;Gene: SS01463

Query Match 75.0%; Score 36; DB 2; Length 417;  
Best Local Similarity 66.7%; Pred. No. 31;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
| : | | | |  
Db 173 VHNFVPLFP 181

#### RESULT 3

S62459

hypothetical protein SPAC2G11.04 - fission yeast (*Schizosaccharomyces pombe*)

C;Species: *Schizosaccharomyces pombe*

C;Date: 16-May-1996 #sequence\_revision 13-Mar-1997 #text\_change 21-Jan-2000

C;Accession: T38570; S62459

R;Badcock, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, October 1995

A;Reference number: Z21745  
A;Accession: T38570  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-301 <BA2>  
A;Cross-references: EMBL:Z54354; NID:g1019398; PIDN:CAA91169.1; PID:g1019402;  
GSPDB:GN00066; SPDB:SPAC2G11.04  
A;Experimental source: strain 972h-; cosmid c2G11  
C;Genetics:  
A;Gene: SPAC2G11.04  
A;Map position: 1L

Query Match 72.9%; Score 35; DB 2; Length 301;  
Best Local Similarity 77.8%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
|| | ||||  
Db 153 VELYDPLFP 161

#### RESULT 4

T37731

probable dihydroorotase - fission yeast (*Schizosaccharomyces pombe*)

C;Species: *Schizosaccharomyces pombe*

C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000

C;Accession: T37731

R;Davis, P.; Churcher, C.M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, October 1999

A;Reference number: Z21741

A;Accession: T37731

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-337 <DAV>

A;Cross-references: EMBL:AL121745; PIDN:CAB57401.1; GSPDB:GN00066;  
SPDB:SPAC16.03c

A;Experimental source: strain 972h-; cosmid c16

C;Genetics:

A;Gene: SPDB:SPAC16.03c

A;Map position: 1

A;Introns: 39/3

C;Superfamily: dihydroorotase

Query Match 72.9%; Score 35; DB 2; Length 337;  
Best Local Similarity 66.7%; Pred. No. 39;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
|||| | :|  
Db 109 VESYEPFYP 117

#### RESULT 5

C86837

amino acid antiporter [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL1403)

C;Species: *Lactococcus lactis* subsp. *lactis*

C;Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C;Accession: C86837  
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich, S.D.; Sorokin, A.  
Genome Res. 11, 731-753, 2001  
A;Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403.  
A;Reference number: A86625; MUID:21235186; PMID:11337471  
A;Accession: C86837  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-466 <STO>  
A;Cross-references: GB:AE005176; PID:g12724715; PIDN:AAK05797.1; GSPDB:GN00146  
A;Experimental source: strain IL1403  
C;Genetics:  
A;Gene: yrfD

Query Match 72.9%; Score 35; DB 2; Length 466;  
Best Local Similarity 71.4%; Pred. No. 55;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
||:|:|  
Db 365 SYIPVFP 371

#### RESULT 6

H70559

hypothetical protein Rv1634 - *Mycobacterium tuberculosis* (strain H37RV)

C;Species: *Mycobacterium tuberculosis*

C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000

C;Accession: H70559

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry III, C.E.; Tekaiia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A;Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence.

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: H70559

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-471 <COL>

A;Cross-references: GB:Z95554; GB:AL123456; NID:g3261771; PIDN:CAB08887.1; PID:g2113918

A;Experimental source: strain H37Rv

C;Genetics:

A;Gene: Rv1634

C;Superfamily: tetracycline resistance protein

Query Match 72.9%; Score 35; DB 2; Length 471;  
Best Local Similarity 75.0%; Pred. No. 55;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VESYVPLF 8  
|::||||  
Db 292 VDTYVPLF 299

RESULT 7

F71565

probable isoleucine-tRNA ligase (EC 6.1.1.5) - Chlamydia trachomatis (serotype D, strain UW3/Cx)

C;Species: Chlamydia trachomatis

C;Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 03-Jun-2002

C;Accession: F71565

R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, W.P.; Olinger, L.; Tatusov, R.L.; Zhao, Q.; Koonin, E.V.; Davis, R.W. Science 282, 754-759, 1998

A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis.

A;Reference number: A71570; MUID:99000809; PMID:9784136

A;Accession: F71565

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1036 <ARN>

A;Cross-references: GB:AE001277; GB:AE001273; NID:g3328408; PIDN:AAC67609.1; PID:g3328409

A;Experimental source: serotype D, strain UW-3/Cx

C;Genetics:

A;Gene: ileS

C;Superfamily: isoleucine-tRNA ligase

C;Keywords: ligase

Query Match 72.9%; Score 35; DB 2; Length 1036;  
Best Local Similarity 75.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9  
: || ||||  
Db 284 QSYEPLFP 291

RESULT 8

T18419

hypothetical protein C0125w - malaria parasite (Plasmodium falciparum)

C;Species: Plasmodium falciparum

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jun-2000

C;Accession: T18419

R;Lawson, D.; Bowman, S.; Barrell, B.  
submitted to the EMBL Data Library, July 1997

A;Reference number: Z18934

A;Accession: T18419

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1365 <LAW>

A;Cross-references: EMBL:Z97348; NID:e1323671; PID:e1323674; PIDN:CAB10570.1

C;Genetics:

A;Note: C0125w



Query Match 72.9%; Score 35; DB 2; Length 1365;  
Best Local Similarity 62.5%; Pred. No. 1.7e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VESYVPLF 8  
|:|:|:|  
Db 183 VDSYIPIF 190

RESULT 9

D83688

hypothetical protein BH0308 [imported] - *Bacillus halodurans* (strain C-125)

C;Species: *Bacillus halodurans*

C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C;Accession: D83688

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hiramata, C.; Nakamura, Y.; Ogasawara, N.; Kuhara, S.; Horikoshi, K.

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*.

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: D83688

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-181 <STO>

A;Cross-references: GB:AP001508; GB:BA000004; NID:g10172890; PIDN:BA04027.1; GSPDB:GN00137

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH0308

Query Match 70.8%; Score 34; DB 2; Length 181;  
Best Local Similarity 62.5%; Pred. No. 32;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9  
: ||||:|  
Db 163 DEYVPLYP 170

RESULT 10

AF2088

hypothetical protein alr2261 [imported] - *Nostoc* sp. (strain PCC 7120)

C;Species: *Nostoc* sp. PCC 7120

A;Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002

C;Accession: AF2088

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, M.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kohara, M.; Matsumoto, M.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena* sp. strain PCC 7120.

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AF2088

A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-202 <KUR>  
A;Cross-references: GB:BA000019; PIDN:BAB73960.1; PID:g17131352; GSPDB:GN00179  
A;Experimental source: strain PCC 7120  
C;Genetics:  
A;Gene: alr2261

Query Match 70.8%; Score 34; DB 2; Length 202;  
Best Local Similarity 83.3%; Pred. No. 36;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 YVPLFP 9  
|:||||  
Db 17 YIPLFP 22

RESULT 11

T10766  
patatin-like latex allergen Hev b3 - Para rubber tree  
N;Alternate names: Hev b 3  
C;Species: Hevea brasiliensis (Para rubber tree)  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Jul-1999  
C;Accession: T10766  
R;Wagner, B.; Krebitz, M.; Scheiner, O.; Breiteneder, H.  
submitted to the EMBL Data Library, January 1998  
A;Description: Molecular characterization and expression of Hev b 3, a major  
latex allergen for spina bifida patients.  
A;Reference number: Z17128  
A;Accession: T10766  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-204 <WAG>  
A;Cross-references: EMBL:AJ223388; NID:e1311128; PID:e1311129  
A;Experimental source: cv. RRIM 600; mature plant; laticifers  
C;Genetics:  
A;Gene: b3

Query Match 70.8%; Score 34; DB 2; Length 204;  
Best Local Similarity 66.7%; Pred. No. 36;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
| ||:| |  
Db 183 VSSYLPLLP 191

RESULT 12

A40359  
aliphatic amidase expression-regulating protein - Pseudomonas aeruginosa  
C;Species: Pseudomonas aeruginosa  
C;Date: 21-Feb-1992 #sequence\_revision 21-Feb-1992 #text\_change 30-Sep-1993  
C;Accession: A40359  
R;Wilson, S.; Drew, R.  
J. Bacteriol. 173, 4914-4921, 1991

A;Title: Cloning and DNA sequence of *amiC*, a new gene regulating expression of the *Pseudomonas aeruginosa* aliphatic amidase, and purification of the *amiC* product.

A;Reference number: A40359; MUID:91317707; PMID:1907262

A;Accession: A40359

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-385 <WIL>

A;Cross-references: GB:X13776

Query Match 70.8%; Score 34; DB 2; Length 385;  
Best Local Similarity 62.5%; Pred. No. 70;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9  
| |::|  
Db 174 EIYIPLYP 181

#### RESULT 13

C83226

aliphatic amidase expression-regulating protein PA3364 [imported] - *Pseudomonas aeruginosa* (strain PA01)

C;Species: *Pseudomonas aeruginosa*

C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C;Accession: C83226

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Brinkman, F.S.L.; Hufnagle, W.O.; Kowalik, D.J.; Lagrou, M.; Garber, R.L.; Goltry, L.; Tolentino, E.; Westbrook-Wadman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, R.M.; Smith, K.A.; Spencer, D.H.; Wong, G.K.S.; Wu, Z.; Paulsen, I.T.; Reizer, J.; Saier, M.H.; Hancock, R.E.W.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen.

A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: C83226

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-385 <STO>

A;Cross-references: GB:AE004758; GB:AE004091; NID:g9949489; PIDN:AAG06752.1;

GSPDB:GN00131; PASP:PA3364

A;Experimental source: strain PA01

C;Genetics:

A;Gene: *amiC*; PA3364

Query Match 70.8%; Score 34; DB 2; Length 385;  
Best Local Similarity 62.5%; Pred. No. 70;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9  
| |::|  
Db 174 EIYIPLYP 181

#### RESULT 14

S66738

hypothetical protein YOL053w - yeast (*Saccharomyces cerevisiae*)  
 N;Alternate names: hypothetical protein O1266; hypothetical protein O1269  
 C;Species: *Saccharomyces cerevisiae*  
 C;Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 19-Apr-2002  
 C;Accession: S66738; S66745; S59297; S59298; S61727; S61728  
 R;Ansorge, W.; Benes, V.; Rechmann, S.; Schwager, C.; Teodoru, C.; Voss, H.;  
 Wiemann, S.  
 submitted to the Protein Sequence Database, July 1996  
 A;Reference number: S66723  
 A;Accession: S66738  
 A;Molecule type: DNA  
 A;Residues: 1-395 <ANS>  
 A;Cross-references: EMBL:Z74795; NID:g1419860; PID:g1419861; MIPS:YOL053w  
 A;Experimental source: strain S288C  
 R;Feldmann, H.; Mannhaupt, G.; Vetter, I.  
 submitted to the Protein Sequence Database, July 1996  
 A;Reference number: S66743  
 A;Accession: S66745  
 A;Molecule type: DNA  
 A;Residues: 1-395 <FEL>  
 A;Cross-references: EMBL:Z74795; NID:g1419860; PID:g1419861; MIPS:YOL053w  
 A;Experimental source: strain S288C  
 R;Mannhaupt, G.; Vetter, I.; Schwarzlose, C.; Mitzel, S.; Feldmann, H.  
 submitted to the EMBL Data Library, August 1995  
 A;Description: Analysis of a 26kb region on the left arm of yeast chromosome XV.  
 A;Reference number: S59285  
 A;Accession: S59297  
 A;Molecule type: DNA  
 A;Residues: 1-128, 'DTVLT' <FEW>  
 A;Cross-references: EMBL:X91067; NID:g984177; PID:g984190  
 A;Accession: S59298  
 A;Molecule type: DNA  
 A;Residues: 228-395 <FEF>  
 A;Cross-references: EMBL:X91067; NID:g984177; PID:g984191  
 R;Mannhaupt, G.; Vetter, I.; Schwarzlose, C.; Mitzel, S.; Feldmann, H.  
 Yeast 12, 67-76, 1996  
 A;Title: Analysis of a 26 kb region on the left arm of yeast chromosome XV.  
 A;Reference number: S61715; MUID:96381248; PMID:8789261  
 A;Accession: S61727  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-128, 'DTVLT' <MAN>  
 A;Cross-references: EMBL:X91067; NID:g984177; PIDN:CAA62534.1; PID:g984190  
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August  
 1995  
 A;Accession: S61728  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 228-395 <MAW>  
 A;Cross-references: EMBL:X91067; NID:g984177; PIDN:CAA62535.1; PID:g984191  
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August  
 1995  
 C;Genetics:  
 A;Cross-references: SGD:S0005414  
 A;Map position: 15L  
 A;Note: YOL053w  
 C;Superfamily: *Saccharomyces cerevisiae* hypothetical protein YOL053w

Query Match 70.8%; Score 34; DB 2; Length 395;  
Best Local Similarity 62.5%; Pred. No. 72;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9  
:|::|  
Db 184 QSFLPLFP 191

RESULT 15

T38188  
probable DNA repair helicase - fission yeast (*Schizosaccharomyces pombe*)  
C;Species: *Schizosaccharomyces pombe*  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C;Accession: T38188  
R;Lye, G.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, October 1995  
A;Reference number: Z21776  
A;Accession: T38188  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-768 <LYE>  
A;Cross-references: EMBL:Z54285; PIDN:CAA91068.2; GSPDB:GN00066;  
SPDB:SPAC22F3.03c  
A;Experimental source: strain 972h-; cosmid c22F3  
C;Genetics:  
A;Gene: SPDB:SPAC22F3.03c  
A;Map position: 1  
A;Introns: 60/2; 93/2; 118/3; 147/2; 211/1; 235/3; 279/3; 301/1; 334/1; 386/3;  
411/1; 616/2; 639/3; 722/3

Query Match 70.8%; Score 34; DB 2; Length 768;  
Best Local Similarity 66.7%; Pred. No. 1.5e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
|| |:  
Db 748 VEDIPLFP 756

Search completed: August 24, 2004, 15:53:01  
Job time : 10.7313 secs

OM protein - protein search, using sw model

Run on: August 24, 2004, 15:51:19 ; Search time 32.5746 Seconds  
(without alignments)  
86.825 Million cell updates/sec

Title: US-09-641-801-31  
Perfect score: 48  
Sequence: 1 VESYVPLFP 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1295152 seqs, 314255058 residues

Total number of hits satisfying chosen parameters: 1295152

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query					
No.	Score	Match	Length	DB	ID	Description

1	48	100.0	9	14	US-10-281-652-31	Sequence 31, Appl
2	36	75.0	190	16	US-10-437-963-170120	Sequence 170120,
3	36	75.0	254	12	US-10-343-357-8	Sequence 8, Appli
4	36	75.0	318	12	US-10-276-774-2663	Sequence 2663, Ap
5	36	75.0	452	12	US-10-282-122A-72546	Sequence 72546, A
6	36	75.0	523	9	US-09-740-029-2	Sequence 2, Appli
7	36	75.0	523	9	US-09-895-728-2	Sequence 2, Appli
8	36	75.0	523	12	US-10-258-080-2	Sequence 2, Appli
9	36	75.0	523	15	US-10-094-749-2927	Sequence 2927, Ap
10	36	75.0	523	16	US-10-381-898-13	Sequence 13, Appl
11	35	72.9	66	16	US-10-437-963-153620	Sequence 153620,
12	35	72.9	97	15	US-10-289-762-71	Sequence 71, Appl
13	35	72.9	152	9	US-09-738-626-5216	Sequence 5216, Ap
14	35	72.9	169	12	US-10-425-114-68398	Sequence 68398, A
15	35	72.9	237	12	US-10-425-114-38324	Sequence 38324, A
16	35	72.9	255	12	US-10-425-114-69055	Sequence 69055, A
17	35	72.9	305	15	US-10-104-047-2670	Sequence 2670, Ap
18	35	72.9	374	16	US-10-437-963-124054	Sequence 124054,
19	35	72.9	411	12	US-10-425-114-71827	Sequence 71827, A
20	35	72.9	446	16	US-10-437-963-178429	Sequence 178429,
21	35	72.9	449	16	US-10-437-963-108868	Sequence 108868,
22	35	72.9	449	16	US-10-437-963-178427	Sequence 178427,
23	35	72.9	458	12	US-10-282-122A-57060	Sequence 57060, A
24	35	72.9	649	12	US-10-425-114-55196	Sequence 55196, A
25	34	70.8	64	10	US-09-764-872-274	Sequence 274, App
26	34	70.8	73	12	US-10-424-599-247680	Sequence 247680,
27	34	70.8	206	12	US-10-282-122A-63019	Sequence 63019, A
28	34	70.8	241	12	US-10-424-599-252357	Sequence 252357,
29	34	70.8	385	12	US-10-389-647-695	Sequence 695, App
30	34	70.8	651	15	US-10-369-493-8731	Sequence 8731, Ap
31	34	70.8	989	15	US-10-435-696-40	Sequence 40, Appl
32	34	70.8	1014	15	US-10-369-493-2077	Sequence 2077, Ap
33	33	68.8	15	10	US-09-880-748-2310	Sequence 2310, Ap
34	33	68.8	15	12	US-10-293-418-2310	Sequence 2310, Ap
35	33	68.8	16	10	US-09-880-748-2263	Sequence 2263, Ap
36	33	68.8	16	12	US-10-293-418-2263	Sequence 2263, Ap
37	33	68.8	40	12	US-10-424-599-234428	Sequence 234428,
38	33	68.8	44	9	US-09-864-761-39037	Sequence 39037, A
39	33	68.8	68	16	US-10-437-963-168174	Sequence 168174,
40	33	68.8	87	12	US-10-424-599-256748	Sequence 256748,
41	33	68.8	172	12	US-10-424-599-155383	Sequence 155383,
42	33	68.8	223	12	US-10-424-599-150707	Sequence 150707,
43	33	68.8	250	10	US-09-880-748-76	Sequence 76, Appl
44	33	68.8	250	12	US-10-293-418-76	Sequence 76, Appl
45	33	68.8	251	10	US-09-880-748-64	Sequence 64, Appl

#### ALIGNMENTS

RESULT 1  
 US-10-281-652-31  
 ; Sequence 31, Application US/10281652  
 ; Publication No. US20030091606A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: STANTON, G. John

```
; APPLICANT: HUGHES, Thomas K.
; APPLICANT: BOLDOGH, Istvan
; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND
; TITLE OF INVENTION: ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS
; FILE REFERENCE: 265.00220101
; CURRENT APPLICATION NUMBER: US/10/281,652
; CURRENT FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: US/09/641,803
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/149,310
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-10-281-652-31
```

```
Query Match          100.0%; Score 48; DB 14; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches      9; Conservative    0; Mismatches    0; Indels      0; Gaps      0;
```

```
Qy      1 VESYVPLFP 9
        |||||
Db      1 VESYVPLFP 9
```

# RESULT 2

US-10-437-963-170120

```
; Sequence 170120, Application US/10437963
```

```
; Publication No. US20040123343A1
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: La Rosa, Thomas J.
```

```
; APPLICANT: Kovalic, David K.
```

```
; APPLICANT: Zhou, Yihua
```

```
; APPLICANT: Cao, Yongwei
```

```
; APPLICANT: Wu, Wei
```

```
; APPLICANT: Boukharov, Andrey A.
```

```
; APPLICANT: Barbazuk, Brad
```

```
; APPLICANT: Li, Ping
```

```
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules
Associated With
```

```
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
```

```
; FILE REFERENCE: 38-21(53221)B
```

```
; CURRENT APPLICATION NUMBER: US/10/437,963
```

```
; CURRENT FILING DATE: 2003-05-14
```

```
; NUMBER OF SEQ ID NOS: 204966
```

```
; SEQ ID NO 170120
```

```
; LENGTH: 190
```

```
; TYPE: PRT
```

```
; ORGANISM: Oryza sativa
```

```
; FEATURE:
```

```
; NAME/KEY: unsure
```



; LOCATION: (1)..(190)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_68475C.1.pep  
US-10-437-963-170120

Query Match 75.0%; Score 36; DB 16; Length 190;  
Best Local Similarity 62.5%; Pred. No. 96;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9  
:|::|  
Db 22 DSFIPLFP 29

RESULT 3

US-10-343-357-8

; Sequence 8, Application US/10343357  
; Publication No. US20040058341A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y.Tom  
; APPLICANT: ELLIOTT, Vicki S.; RAMKUMAR, Jayalaxmi  
; APPLICANT: YAO, Monique G.; BURFORD, Neil  
; APPLICANT: WANG, Yumei E.; STEWART, Elizabeth A.  
; APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.  
; APPLICANT: LEE, Ernestine A.; HAFALIA, April J.A.  
; APPLICANT: LU, Dyung Aina M.; TRIBOULEY, Catherine M.  
; APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Mariah R.  
; APPLICANT: YUE, Henry; WARREN, Bridget A.  
; APPLICANT: NGUYEN, Danniel B.; CHAWLA, Narinder K.  
; APPLICANT: KEARNEY, Liam  
; TITLE OF INVENTION: PROTEIN PHOSPHATASES  
; FILE REFERENCE: PI-0173 PCT  
; CURRENT APPLICATION NUMBER: US/10/343,357  
; CURRENT FILING DATE: 2003-01-28  
; PRIOR APPLICATION NUMBER: PCT/US01/23716  
; PRIOR FILING DATE: 2001-07-26  
; PRIOR APPLICATION NUMBER: US 60/221,679  
; PRIOR FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: US 60/223,272  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/224,309  
; PRIOR FILING DATE: 2000-08-10  
; PRIOR APPLICATION NUMBER: US 60/226,728  
; PRIOR FILING DATE: 2000-08-18  
; PRIOR APPLICATION NUMBER: US 60/229,254  
; PRIOR FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: US 60/231,366  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PERL Program  
; SEQ ID NO 8  
; LENGTH: 254  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature

; OTHER INFORMATION: Incyte ID No. US20040058341A1 4555838CD1  
US-10-343-357-8

Query Match 75.0%; Score 36; DB 12; Length 254;  
Best Local Similarity 71.4%; Pred. No. 1.3e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
||:|:|  
Db 224 SYIPIFP 230

RESULT 4

US-10-276-774-2663  
; Sequence 2663, Application US/10276774  
; Publication No. US20040053245A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; APPLICANT: Tang, Y, Tom et al  
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-030  
; CURRENT APPLICATION NUMBER: US/10/276,774  
; CURRENT FILING DATE: 2002-11-18  
; PRIOR APPLICATION NUMBER: 09/560,875  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 09/496,914  
; PRIOR FILING DATE: 2000-02-03  
; NUMBER OF SEQ ID NOS: 2700  
; SOFTWARE: Custom  
; SEQ ID NO 2663  
; LENGTH: 318  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-276-774-2663

Query Match 75.0%; Score 36; DB 12; Length 318;  
Best Local Similarity 85.7%; Pred. No. 1.6e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
||||:|  
Db 207 SYVPVFP 213

RESULT 5

US-10-282-122A-72546  
; Sequence 72546, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John

```

; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72546
;   LENGTH: 452
;   TYPE: PRT
;   ORGANISM: Streptococcus mutans
US-10-282-122A-72546

```

```

Query Match          75.0%;   Score 36;   DB 12;   Length 452;
Best Local Similarity 85.7%;   Pred. No. 2.4e+02;
Matches      6;   Conservative    1;   Mismatches    0;   Indels      0;   Gaps      0;

```

```

Qy          3 SYVPLFP 9
           ||||:|
Db         364 SYVPVFP 370

```

# RESULT 6

US-09-740-029-2

```

; Sequence 2, Application US/09740029
; Patent No. US20020082194A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000928

```

; CURRENT APPLICATION NUMBER: US/09/740,029  
; CURRENT FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 523  
; TYPE: PRT  
; ORGANISM: Human  
US-09-740-029-2

Query Match 75.0%; Score 36; DB 9; Length 523;  
Best Local Similarity 85.7%; Pred. No. 2.8e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
| | | | : | |  
Db 182 SYVPVFP 188

RESULT 7

US-09-895-728-2  
; Sequence 2, Application US/09895728  
; Patent No. US20020123475A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Leiby, Kevin  
; APPLICANT: Spaltmann, Frank  
; APPLICANT: Cook, William  
; TITLE OF INVENTION: 32626, A NOVEL HUMAN  
; TITLE OF INVENTION: UDP-GLYCOSYLTRANSFERASE AND USES THEREOF  
; FILE REFERENCE: 38155-20018.00  
; CURRENT APPLICATION NUMBER: US/09/895,728  
; CURRENT FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/215,749  
; PRIOR FILING DATE: 2000-06-30  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 523  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-895-728-2

Query Match 75.0%; Score 36; DB 9; Length 523;  
Best Local Similarity 85.7%; Pred. No. 2.8e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
| | | | : | |  
Db 182 SYVPVFP 188

RESULT 8

US-10-258-080-2  
; Sequence 2, Application US/10258080  
; Publication No. US20040029125A1  
; GENERAL INFORMATION:

```

; APPLICANT: Incyte Genomics, Inc.
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: BURFORD, Neil
; APPLICANT: RING, Huijun Z.
; APPLICANT: LAL, Preeti G.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: YAO, Monique G.
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: DAS, Debopriya
; APPLICANT: SANJANWALA, Madhusudan M.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: REDDY, Roopa M.
; APPLICANT: KHAN, Farrah A.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: AU-YOUNG, Janice K.
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0070 USN
; CURRENT APPLICATION NUMBER: US/10/258,080
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/203,509
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/202,234
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/200,185
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/198,403
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US01/11869
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US 60/197,590
; PRIOR FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040029125A1 2516747CD1
US-10-258-080-2

```

```

Query Match          75.0%; Score 36; DB 12; Length 523;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches      6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      3 SYVPLFP 9
        ||||:|
Db      182 SYVPVFP 188

```

RESULT 9

US-10-094-749-2927  
; Sequence 2927, Application US/10094749  
; Publication No. US20030219741A1  
; GENERAL INFORMATION:  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: OTSUKI, TETSUJI  
; APPLICANT: WAKAMATSU, AI  
; APPLICANT: SATO, HIROYUKI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: YAMAMOTO, JUN-ICHI  
; APPLICANT: ISONO, YUUKO  
; APPLICANT: HIO, YURI  
; APPLICANT: OTSUKA, KAORU  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: IRIE, RYOTARO  
; APPLICANT: TAMECHIKA, ICHIRO  
; APPLICANT: SEKI, NAOHIKO  
; APPLICANT: YOSHIKAWA, TSUTOMU  
; APPLICANT: OTSUKA, MOTOYUKI  
; APPLICANT: NAGAHARI, KENJI  
; APPLICANT: MASUHO, YASUHIKO  
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA  
; FILE REFERENCE: 084335/0160  
; CURRENT APPLICATION NUMBER: US/10/094,749  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/350,435  
; PRIOR FILING DATE: 2002-01-24  
; PRIOR APPLICATION NUMBER: JP 2001-328381  
; PRIOR FILING DATE: 2001-09-14  
; NUMBER OF SEQ ID NOS: 3381  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2927  
; LENGTH: 523  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-094-749-2927

Query Match 75.0%; Score 36; DB 15; Length 523;  
Best Local Similarity 85.7%; Pred. No. 2.8e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
| | | | : | |  
Db 182 SYVPVFP 188

RESULT 10  
US-10-381-898-13  
; Sequence 13, Application US/10381898  
; Publication No. US20040086887A1  
; GENERAL INFORMATION:  
; APPLICANT: AZIMZAI, Yalda; BAUGHN, Mariah R.;  
; APPLICANT: BOROWSKY, Mark L.; DING, Li;  
; APPLICANT: DUGGAN, Brendan; ELLIOTT, Vicki S.;  
; APPLICANT: GANDHI, Ameena R.; GRIFFIN, Jennifer A.;  
; APPLICANT: HAFALIA, April J.A.; ISON, Craig H.;

```

; APPLICANT: KHAN, Farrah A.; LAL, Preeti G.;
; APPLICANT: LEE, Ernestine A.; LU, Dyung Aina M.;
; APPLICANT: ARVIZU, Chandra S.; POLICKY, Jennifer L.;
; APPLICANT: RAMKUMAR, Jayalaxmi; RING, Huijun Z.;
; APPLICANT: SANJANWALA, Madhusudan M.;
; APPLICANT: TANG, Y. Tom; TRIBOULEY, Catherine M.;
; APPLICANT: CHAWLA, Narinder K.; WALSH, Roderick T.;
; APPLICANT: WARREN, Bridget; XU, Yuming;
; APPLICANT: YANG, Junming; YAO, Monique; YUE, Henry
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0233 USN
; CURRENT APPLICATION NUMBER: US/10/381,898
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/US01/30662
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/236,947
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/238,864
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/242,323
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/247,581
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/249,519
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/252,834
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/250,567
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7484241CD1
US-10-381-898-13

```

```

Query Match          75.0%; Score 36; DB 16; Length 523;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches      6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      3 SYVPLFP 9
        ||||:|
Db     182 SYVPVFP 188

```

```

RESULT 11
US-10-437-963-153620
; Sequence 153620, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua

```

; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Wu, Wei  
 ; APPLICANT: Boukharov, Andrey A.  
 ; APPLICANT: Barbazuk, Brad  
 ; APPLICANT: Li, Ping  
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules  
 Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53221)B  
 ; CURRENT APPLICATION NUMBER: US/10/437,963  
 ; CURRENT FILING DATE: 2003-05-14  
 ; NUMBER OF SEQ ID NOS: 204966  
 ; SEQ ID NO 153620  
 ; LENGTH: 66  
 ; TYPE: PRT  
 ; ORGANISM: Oryza sativa  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_53558C.1.pep  
 US-10-437-963-153620

Query Match 72.9%; Score 35; DB 16; Length 66;  
 Best Local Similarity 85.7%; Pred. No. 49;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
 |:|||||  
 Db 36 SFVPLFP 42

# RESULT 12

US-10-289-762-71

; Sequence 71, Application US/10289762  
 ; Publication No. US20040006218A1  
 ; GENERAL INFORMATION:

; APPLICANT: Griffais, R.

; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides,  
 fragments

; TITLE OF INVENTION: thereof and uses thereof, in particular for the  
 diagnosis, prevention

; TITLE OF INVENTION: and treatment of infection

; FILE REFERENCE: 9710-003-999

; CURRENT APPLICATION NUMBER: US/10/289,762

; CURRENT FILING DATE: 2003-03-27

; NUMBER OF SEQ ID NOS: 6849

; SEQ ID NO 71

; LENGTH: 97

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-10-289-762-71

Query Match 72.9%; Score 35; DB 15; Length 97;  
 Best Local Similarity 75.0%; Pred. No. 73;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9  
 |:|||||  
 Db 1 ETVVPLLP 8



RESULT 13

US-09-738-626-5216

; Sequence 5216, Application US/09738626  
 ; Publication No. US20020197605A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NAKAGAWA, SATOSHI  
 ; APPLICANT: MIZOGUCHI, HIROSHI  
 ; APPLICANT: ANDO, SEIKO  
 ; APPLICANT: HAYASHI, MIKIRO  
 ; APPLICANT: OCHIAI, KEIKO  
 ; APPLICANT: YOKOI, HARUHIKO  
 ; APPLICANT: TATEISHI, NAOKO  
 ; APPLICANT: SENOH, AKIHIRO  
 ; APPLICANT: IKEDA, MASATO  
 ; APPLICANT: OZAKI, AKIO  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-125  
 ; CURRENT APPLICATION NUMBER: US/09/738,626  
 ; CURRENT FILING DATE: 2000-12-18  
 ; PRIOR APPLICATION NUMBER: JP 99/377484  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: JP 00/159162  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; PRIOR FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 7059  
 ; SOFTWARE: PatentIn ver. 3.0  
 ; SEQ ID NO 5216  
 ; LENGTH: 152  
 ; TYPE: PRT  
 ; ORGANISM: Corynebacterium glutamicum  
 US-09-738-626-5216

Query Match 72.9%; Score 35; DB 9; Length 152;  
 Best Local Similarity 55.6%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
 :::||| ||  
 Db 116 LDTYVPFFP 124

RESULT 14

US-10-425-114-68398

; Sequence 68398, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated  
 With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 68398  
; LENGTH: 169  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMROB73004D12\_FLI.pep  
US-10-425-114-68398

Query Match 72.9%; Score 35; DB 12; Length 169;  
Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
|:|||||  
Db 120 SFVPLFP 126

RESULT 15

US-10-425-114-38324  
; Sequence 38324, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated  
With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 38324  
; LENGTH: 237  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3242-151-D11\_FLI.pep  
US-10-425-114-38324

Query Match 72.9%; Score 35; DB 12; Length 237;  
Best Local Similarity 87.5%; Pred. No. 1.9e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESYVPLF 8  
| |||||  
Db 6 VASYVPLF 13

Search completed: August 24, 2004, 16:41:29  
Job time : 35.5746 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 15:23:00 ; Search time 27.806 Seconds  
(without alignments)  
102.124 Million cell updates/sec

Title: US-09-641-801-31  
Perfect score: 48  
Sequence: 1 VESYVPLFP 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriaphage:\*  
17: sp\_archaeophages:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	
No.	Score Match Length DB ID	Description
-----		

1	41	85.4	293	10	Q9C9J2	Q9c9j2 arabidopsis
2	40	83.3	290	16	Q7U931	Q7u931 synechococc
3	40	83.3	1087	5	Q9BMG3	Q9bmg3 leishmania
4	38	79.2	239	13	Q803P3	Q803p3 brachydanio
5	38	79.2	248	11	Q9D6P7	Q9d6p7 mus musculu
6	38	79.2	254	11	Q9D3B1	Q9d3b1 mus musculu
7	38	79.2	258	11	Q9JLK1	Q9jlk1 mus musculu
8	38	79.2	281	11	Q9QY80	Q9qy80 mus musculu
9	38	79.2	288	4	Q96FW7	Q96fw7 homo sapien
10	38	79.2	288	4	Q9UHX2	Q9uhx2 homo sapien
11	38	79.2	288	6	Q9N1R5	Q9n1r5 ovis aries
12	38	79.2	290	4	Q9HB93	Q9hb93 homo sapien
13	36	75.0	252	4	Q8IYS9	Q8iys9 homo sapien
14	36	75.0	403	16	Q7VAR5	Q7var5 prochloroco
15	36	75.0	417	17	Q97Y76	Q97y76 sulfolobus
16	36	75.0	452	16	Q8DW18	Q8dw18 streptococc
17	36	75.0	476	16	Q8DK24	Q8dk24 synechococc
18	36	75.0	483	16	Q88HD9	Q88hd9 pseudomonas
19	36	75.0	497	16	Q88HE0	Q88he0 pseudomonas
20	36	75.0	523	4	Q96DM6	Q96dm6 homo sapien
21	35	72.9	152	16	Q8NQ85	Q8nq85 corynebacte
22	35	72.9	192	4	O00504	O00504 homo sapien
23	35	72.9	219	16	Q89BI7	Q89bi7 bradyrhizob
24	35	72.9	246	16	Q8DWB1	Q8dwb1 streptococc
25	35	72.9	305	4	Q8NB02	Q8nb02 homo sapien
26	35	72.9	328	16	Q8G0C1	Q8g0c1 brucella su
27	35	72.9	332	16	Q98KU6	Q98ku6 rhizobium l
28	35	72.9	430	10	Q7XPK8	Q7xpk8 oryza sativ
29	35	72.9	439	10	Q84QG6	Q84qg6 hordeum vul
30	35	72.9	449	10	Q7XPK7	Q7xpk7 oryza sativ
31	35	72.9	458	16	Q837U6	Q837u6 enterococcu
32	35	72.9	466	16	Q9CEY5	Q9cey5 lactococcus
33	35	72.9	468	16	Q8ELP9	Q8elp9 oceanobacil
34	35	72.9	471	16	O06151	O06151 mycobacteri
35	35	72.9	471	16	Q7VEV8	Q7vev8 mycobacteri
36	35	72.9	475	17	Q8TK26	Q8tk26 methanosarc
37	35	72.9	849	16	Q8RDR3	Q8rdr3 fusobacteri
38	35	72.9	956	11	Q8R004	Q8r004 mus musculu
39	35	72.9	956	11	Q9WVF1	Q9wvf1 mus musculu
40	35	72.9	987	11	Q99K74	Q99k74 mus musculu
41	35	72.9	1365	5	O77308	O77308 plasmodium
42	34	70.8	181	16	Q9KG07	Q9kg07 bacillus ha
43	34	70.8	182	17	Q8ZZB3	Q8zzb3 pyrobaculum
44	34	70.8	202	16	Q8YUS3	Q8yus3 anabaena sp
45	34	70.8	241	4	Q9H5V2	Q9h5v2 homo sapien

# ALIGNMENTS

## RESULT 1

Q9C9J2

ID Q9C9J2 PRELIMINARY; PRT; 293 AA.  
AC Q9C9J2;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Putative carboxypeptidase, 8589-6746.  
GN F26A9.4.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,  
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,  
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;  
RT "Arabidopsis thaliana chromosome 1 BAC F26A9 genomic sequence."  
RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AC016163; AAG51831.1; -.  
DR GO; GO:0004182; F:carboxypeptidase A activity; IEA.  
DR GO; GO:0004180; F:carboxypeptidase activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR000834; Peptidase\_M14.  
DR InterPro; IPR008575; Peptidase\_M14B.  
DR Pfam; PF05885; DUF857; 1.  
DR Pfam; PF00246; Zn\_carbOpept; 1.  
DR PRINTS; PR00765; CRBOXYPTASEA.  
DR SMART; SM00631; Zn\_pept; 1.  
DR PROSITE; PS00132; CARBOXYPEPT\_ZN\_1; 1.  
DR PROSITE; PS00133; CARBOXYPEPT\_ZN\_2; 1.  
KW Carboxypeptidase.  
SQ SEQUENCE 293 AA; 33939 MW; 447FE7423F73801F CRC64;

Query Match 85.4%; Score 41; DB 10; Length 293;  
Best Local Similarity 66.7%; Pred. No. 6.2;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
|:|:|:|:  
Db 275 VKSYIPIFP 283

## RESULT 2

Q7U931

ID Q7U931 PRELIMINARY; PRT; 290 AA.  
AC Q7U931;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical.  
GN SYNW0428.  
OS Synechococcus sp. (strain WH8102).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.  
OX NCBI\_TaxID=84588;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22825697; PubMed=12917641;  
RA Palenik B., Brahamsha B., Larimer F.W., Land M., Hauser L., Chain P.,  
RA Lamerdin J., Regala W., Allen E.E., McCarren J., Paulsen I.,  
RA Dufresne A., Partensky F., Webb E.A., Waterbury J.;

RT "The genome of a motile marine Synechococcus.";  
RL Nature 424:1037-1042(2003).  
DR EMBL; BX569690; CAE06943.1; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 290 AA; 33290 MW; 688A9DC2B61CB7AF CRC64;

Query Match 83.3%; Score 40; DB 16; Length 290;  
Best Local Similarity 66.7%; Pred. No. 9.7;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
:| |||:|  
Db 112 IEGYVPVFP 120

### RESULT 3

#### Q9BMG3

ID Q9BMG3 PRELIMINARY; PRT; 1087 AA.  
AC Q9BMG3;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Isoleucine tRNA synthetase (EC 6.1.1.5).  
GN ILERS.  
OS Leishmania donovani.  
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
OX NCBI\_TaxID=5661;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Abdulghani J., Selenke L., Crowe K., Mishra M., Lewis C.S.,  
RA Chaudhuri G.;  
RT "Cloning and characterization of the isoleucine tRNA synthetase gene  
RT from Leishmania donovani.";  
RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF326935; AAG49529.2; -.  
DR HSSP; P56690; 1ILE.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004822; F:isoleucine-tRNA ligase activity; IEA.  
DR GO; GO:0016874; F:ligase activity; IEA.  
DR GO; GO:0006428; P:isoleucyl-tRNA aminoacylation; IEA.  
DR InterPro; IPR002300; tRNA-synt\_1a.  
DR InterPro; IPR001412; tRNA-synt\_I.  
DR InterPro; IPR002301; tRNA-synt\_ile.  
DR InterPro; IPR009008; ValRS\_IleRS\_edit.  
DR Pfam; PF00133; tRNA-synt\_1; 1.  
DR PRINTS; PR00984; TRNASYNTHILE.  
DR TIGRFAMs; TIGR00392; ileS; 1.  
DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_I; 1.  
KW Aminoacyl-tRNA synthetase; Ligase.  
SQ SEQUENCE 1087 AA; 124462 MW; 11DD6AB2409146B8 CRC64;

Query Match 83.3%; Score 40; DB 5; Length 1087;  
Best Local Similarity 87.5%; Pred. No. 39;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9  
| |||||

RESULT 4

Q803P3

ID Q803P3 PRELIMINARY; PRT; 239 AA.  
 AC Q803P3;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Similar to RIKEN cDNA 6330408J20 gene.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; BC044396; AAH44396.1; -.  
 DR GO; GO:0004721; F:protein phosphatase activity; IEA.  
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.  
 DR InterPro; IPR007482; PTPLA.  
 DR InterPro; IPR000387; TYR\_phosphatase.  
 DR Pfam; PF04387; PTPLA; 1.  
 DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
 SQ SEQUENCE 239 AA; 27358 MW; AA059EF1316B5078 CRC64;

Query Match 79.2%; Score 38; DB 13; Length 239;  
 Best Local Similarity 85.7%; Pred. No. 20;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
 ||:||||  
 Db 209 SYIPLFP 215

RESULT 5

Q9D6P7

ID Q9D6P7 PRELIMINARY; PRT; 248 AA.  
 AC Q9D6P7;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Protein tyrosine phosphatase-like (Proline instead of catalytic  
 DE arginine), member a.  
 GN PTPLA.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Tongue;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,



RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK010119; BAB26713.1; -.  
 DR MGD; MGI:1353592; Ptpla.  
 DR GO; GO:0004721; F:protein phosphatase activity; IEA.  
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.  
 DR InterPro; IPR007482; PTPLA.  
 DR InterPro; IPR000387; TYR\_phosphatase.  
 DR Pfam; PF04387; PTPLA; 1.  
 DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
 SQ SEQUENCE 248 AA; 28523 MW; C863FBE98127EE0D CRC64;

Query Match 79.2%; Score 38; DB 11; Length 248;  
 Best Local Similarity 85.7%; Pred. No. 21;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 SYVPLFP 9  
 ||:||||  
 Db 218 SYIPLFP 224

# RESULT 6

Q9D3B1

ID Q9D3B1 PRELIMINARY; PRT; 254 AA.  
 AC Q9D3B1;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE 6330408J20Rik protein (RIKEN cDNA 6330408J20 gene).  
 GN 6330408J20RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection";  
 RL Nature 409:685-690(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AK018144; BAB31092.1; -.  
 DR EMBL; BC027289; AAH27289.1; -.  
 DR MGD; MGI:1918007; 6330408J20Rik.  
 DR GO; GO:0004721; F:protein phosphatase activity; IEA.  
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.  
 DR InterPro; IPR007482; PTPLA.  
 DR InterPro; IPR000387; TYR\_phosphatase.  
 DR Pfam; PF04387; PTPLA; 1.  
 DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
 SQ SEQUENCE 254 AA; 28402 MW; 7DB21BB52473EB0C CRC64;

Query Match 79.2%; Score 38; DB 11; Length 254;  
 Best Local Similarity 85.7%; Pred. No. 21;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
 ||:||||  
 Db 224 SYIPLFP 230

#### RESULT 7

Q9JLK1

ID Q9JLK1 PRELIMINARY; PRT; 258 AA.  
 AC Q9JLK1;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Protein tyrosine phosphatase-like protein PTPLB.  
 GN 6330408J20RIK OR PTPLB.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=20112754; PubMed=10644438;  
 RA Uwanogho D.A., Hardcastle Z., Balogh P., Mirza G., Thornburg K.L.,  
 RA Ragoussis J., Sharpe P.T.;  
 RT "Molecular cloning, chromosomal mapping, and developmental expression  
 RT of a novel protein tyrosine phosphatase-like gene."  
 RL Genomics 62:406-416(1999).  
 DR EMBL; AF169286; AAF29489.1; -.  
 DR MGD; MGI:1918007; 6330408J20Rik.  
 DR GO; GO:0004721; F:protein phosphatase activity; IEA.  
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.  
 DR InterPro; IPR007482; PTPLA.  
 DR InterPro; IPR000387; TYR\_phosphatase.  
 DR Pfam; PF04387; PTPLA; 1.  
 DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
 SQ SEQUENCE 258 AA; 28843 MW; 6B46680875346987 CRC64;

Query Match 79.2%; Score 38; DB 11; Length 258;  
 Best Local Similarity 85.7%; Pred. No. 22;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
 ||:||||  
 Db 228 SYIPLFP 234

#### RESULT 8

Q9QY80

ID Q9QY80 PRELIMINARY; PRT; 281 AA.  
 AC Q9QY80;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Putative tyrosine phosphatase.  
 GN PTPLA.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20112754; PubMed=10644438;  
 RA Uwanogho D.A., Hardcastle Z., Balogh P., Mirza G., Thornburg K.L.,  
 RA Ragoussis J., Sharpe P.T.;  
 RT "Molecular cloning, chromosomal mapping, and developmental expression  
 RT of a novel protein tyrosine phosphatase-like gene."  
 RL Genomics 62:406-416(1999).  
 DR EMBL; AF114493; AAF21975.1; -.  
 DR MGD; MGI:1353592; Ptpla.  
 DR GO; GO:0004721; F:protein phosphatase activity; IEA.  
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.  
 DR InterPro; IPR007482; PTPLA.  
 DR InterPro; IPR000387; TYR\_phosphatase.  
 DR Pfam; PF04387; PTPLA; 1.  
 DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
 SQ SEQUENCE 281 AA; 32327 MW; 24D64726C427BB08 CRC64;

Query Match 79.2%; Score 38; DB 11; Length 281;

Best Local Similarity 85.7%; Pred. No. 24;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
||:||||  
Db 251 SYIPLFP 257

RESULT 9

Q96FW7

ID Q96FW7 PRELIMINARY; PRT; 288 AA.  
AC Q96FW7;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Similar to protein tyrosine phosphatase-like (Proline instead of  
DE catalytic arginine), member a.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skeletal muscle;  
RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; BC010353; AAH10353.1; -.  
DR GO; GO:0004721; F:protein phosphatase activity; IEA.  
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.  
DR InterPro; IPR001844; Chaprin\_Cpn60.  
DR InterPro; IPR007482; PTPLA.  
DR InterPro; IPR000387; TYR\_phosphatase.  
DR Pfam; PF04387; PTPLA; 1.  
DR PROSITE; PS00296; CHAPERONINS\_CPN60; 1.  
DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
SQ SEQUENCE 288 AA; 32435 MW; 3926B89706AD3610 CRC64;

Query Match 79.2%; Score 38; DB 4; Length 288;  
Best Local Similarity 85.7%; Pred. No. 24;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
||:||||  
Db 258 SYIPLFP 264

RESULT 10

Q9UHX2

ID Q9UHX2 PRELIMINARY; PRT; 288 AA.  
AC Q9UHX2;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Putative tyrosine phosphatase.  
GN PTPLA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20112754; PubMed=10644438;  
 RA Uwanogho D.A., Hardcastle Z., Balogh P., Mirza G., Thornburg K.L.,  
 RA Ragoussis J., Sharpe P.T.;  
 RT "Molecular cloning, chromosomal mapping, and developmental expression  
 RT of a novel protein tyrosine phosphatase-like gene."  
 RL Genomics 62:406-416(1999).  
 DR EMBL; AF114494; AAF21976.1; -.  
 DR GO; GO:0004725; F:protein tyrosine phosphatase activity; TAS.  
 DR GO; GO:0007275; P:development; TAS.  
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.  
 DR InterPro; IPR001844; Chaprin\_Cpn60.  
 DR InterPro; IPR007482; PTPLA.  
 DR InterPro; IPR000387; TYR\_phosphatase.  
 DR Pfam; PF04387; PTPLA; 1.  
 DR PROSITE; PS00296; CHAPERONINS\_CPN60; 1.  
 DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
 SQ SEQUENCE 288 AA; 32426 MW; 7A6A4274653CC0D9 CRC64;

Query Match 79.2%; Score 38; DB 4; Length 288;  
 Best Local Similarity 85.7%; Pred. No. 24;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
 ||:||||  
 Db 258 SYIPLFP 264

# RESULT 11

Q9N1R5

ID Q9N1R5 PRELIMINARY; PRT; 288 AA.  
 AC Q9N1R5;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Protein tyrosine phosphatase-like protein PTPLA.  
 GN PTPLA.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20112754; PubMed=10644438;  
 RA Uwanogho D.A., Hardcastle Z., Balogh P., Mirza G., Thornburg K.L.,  
 RA Ragoussis J., Sharpe P.T.;  
 RT "Molecular cloning, chromosomal mapping, and developmental expression  
 RT of a novel protein tyrosine phosphatase-like gene."  
 RL Genomics 62:406-416(1999).  
 DR EMBL; AF162707; AAF29469.1; -.  
 DR GO; GO:0004721; F:protein phosphatase activity; IEA.  
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.  
 DR InterPro; IPR001844; Chaprin\_Cpn60.

DR InterPro; IPR007482; PTPLA.  
 DR InterPro; IPR000387; TYR\_phosphatase.  
 DR Pfam; PF04387; PTPLA; 1.  
 DR PROSITE; PS00296; CHAPERONINS\_CPN60; 1.  
 DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
 SQ SEQUENCE 288 AA; 32151 MW; 63B0D84D7B79EF7A CRC64;

Query Match 79.2%; Score 38; DB 6; Length 288;  
 Best Local Similarity 85.7%; Pred. No. 24;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
 ||:||||  
 Db 258 SYIPLFP 264

# RESULT 12

Q9HB93

ID Q9HB93 PRELIMINARY; PRT; 290 AA.  
 AC Q9HB93;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE PTPLA.  
 GN PTPLA.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Li D., Bachinski L., Roberts R.;  
 RT "Expression profile, genomic structure and mutational analysis of  
 RT PTPLA in arrhythmogenic right ventricular dysplasia mapped to  
 RT chromosome 10p12-14.";  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF266852; AAG10713.1; -.  
 DR EMBL; AF266848; AAG10713.1; JOINED.  
 DR EMBL; AF266849; AAG10713.1; JOINED.  
 DR EMBL; AF266850; AAG10713.1; JOINED.  
 DR EMBL; AF266851; AAG10713.1; JOINED.  
 DR Genew; HGNC:9639; PTPLA.  
 DR GO; GO:0004721; F:protein phosphatase activity; IEA.  
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.  
 DR InterPro; IPR001844; Chaprin\_Cpn60.  
 DR InterPro; IPR007482; PTPLA.  
 DR InterPro; IPR000387; TYR\_phosphatase.  
 DR Pfam; PF04387; PTPLA; 1.  
 DR PROSITE; PS00296; CHAPERONINS\_CPN60; 1.  
 DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
 SQ SEQUENCE 290 AA; 32729 MW; 40BBB5D7EDF3654D CRC64;

Query Match 79.2%; Score 38; DB 4; Length 290;  
 Best Local Similarity 85.7%; Pred. No. 24;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9

Db                    ||:||||  
260 SYIPLFP 266

RESULT 13

Q8IYS9

ID    Q8IYS9            PRELIMINARY;            PRT;    252 AA.  
AC    Q8IYS9;  
DT    01-MAR-2003 (TrEMBLrel. 23, Created)  
DT    01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT    01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE    Hypothetical protein.  
OS    Homo sapiens (Human).  
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX    NCBI\_TaxID=9606;  
RN    [1]  
RP    SEQUENCE FROM N.A.  
RC    TISSUE=Testis;  
RA    Strausberg R.;  
RL    Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.  
DR    EMBL; BC035012; AAH35012.1; -.  
DR    GO; GO:0016758; F:transferase activity, transferring hexosyl . . .; IEA.  
DR    GO; GO:0008152; P:metabolism; IEA.  
DR    InterPro; IPR002213; UDP\_gluco\_trans.  
DR    Pfam; PF00201; UDPGT; 1.  
KW    Hypothetical protein.  
SQ    SEQUENCE    252 AA;    29155 MW;    956CC9F9718B8982 CRC64;

Query Match                    75.0%;    Score 36;    DB 4;    Length 252;  
Best Local Similarity    85.7%;    Pred. No. 53;  
Matches       6;    Conservative    1;    Mismatches       0;    Indels       0;    Gaps       0;

Qy                    3 SYVPLFP 9  
                      ||||:||  
Db                    128 SYVPVFP 134

RESULT 14

Q7VAR5

ID    Q7VAR5            PRELIMINARY;            PRT;    403 AA.  
AC    Q7VAR5;  
DT    01-OCT-2003 (TrEMBLrel. 25, Created)  
DT    01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT    01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE    Glycosyltransferase.  
GN    RFAG OR PRO1392.  
OS    Prochlorococcus marinus.  
OC    Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
OC    Prochlorococcus.  
OX    NCBI\_TaxID=1219;  
RN    [1]  
RP    SEQUENCE FROM N.A.  
RC    STRAIN=SARG / CCMP 1375 / SS120;  
RX    MEDLINE=22810154; PubMed=12917486;  
RA    Dufresne A., Salanoubat M., Partensky F., Artiguenave F., Axmann I.M.,  
RA    Barbe V., Duprat S., Galperin M.Y., Koonin E.V., Le Gall F.,

RA Makarova K.S., Ostrowski M., Oztas S., Robert C., Rogozin I.B.,  
 RA Scanlan D.J., Tandeau de Marsac N., Weissenbach J., Wincker P.,  
 RA Wolf Y.I., Hess W.R.;  
 RT "Genome sequence of the cyanobacterium *Prochlorococcus marinus* SS120,  
 RT a nearly minimal oxyphototrophic genome."  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025(2003).  
 DR EMBL; AE017165; AAQ00436.1; -.  
 KW Transferase; Complete proteome.  
 SQ SEQUENCE 403 AA; 45573 MW; A09748E16820DFE8 CRC64;

Query Match 75.0%; Score 36; DB 16; Length 403;  
 Best Local Similarity 66.7%; Pred. No. 86;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
 |:|:|:| |  
 Db 78 VQSYLPLLP 86

# RESULT 15

Q97Y76

ID Q97Y76 PRELIMINARY; PRT; 417 AA.  
 AC Q97Y76;  
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Amino acid transporter.  
 GN SS01463.  
 OS *Sulfolobus solfataricus*.  
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 OC *Sulfolobus*.  
 OX NCBI\_TaxID=2287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;  
 RX MEDLINE=21332296; PubMed=11427726;  
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,  
 RA Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,  
 RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,  
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,  
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,  
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,  
 RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;  
 RT "The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).  
 DR EMBL; AE006762; AAK41691.1; -.  
 DR PIR; D90304; D90304.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.  
 DR GO; GO:0006865; P:amino acid transport; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR002293; AA/rel\_permease1.  
 DR InterPro; IPR002422; AA/rel\_permease2.  
 DR InterPro; IPR004841; Permease\_region.  
 DR Pfam; PF00324; aa\_permeases; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 417 AA; 45077 MW; CBF8D62B18551502 CRC64;



Query Match 75.0%; Score 36; DB 17; Length 417;  
Best Local Similarity 66.7%; Pred. No. 89;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
| ::|||||  
Db 173 VHNEVPLFP 181

Search completed: August 24, 2004, 15:50:58  
Job time : 30.806 secs

OM protein - protein search, using sw model

Run on: August 24, 2004, 14:57:04 ; Search time 4.83582 Seconds  
 (without alignments)  
 96.908 Million cell updates/sec

Title: US-09-641-801-31  
 Perfect score: 48  
 Sequence: 1 VESYVPLFP 9

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	37	77.1	2136	1	YCF2_MARPO	P09975 marchantia
2	35	72.9	301	1	YAB4_SCHPO	Q09806 schizosacch
3	35	72.9	337	1	PYRC_SCHPO	Q9uti0 schizosacch
4	35	72.9	1036	1	SYI_CHLTR	O84022 chlamydia t
5	35	72.9	1194	1	SA2_XENLA	Q9dgn0 xenopus lae
6	34	70.8	204	1	SRPP_HEVBR	O82803 hevea brasi
7	34	70.8	304	1	OTC_PYRAE	Q8zw40 pyrobaculum
8	34	70.8	384	1	AMIC_PSEAE	P27017 pseudomonas
9	34	70.8	395	1	EFTU_FLAFE	P42476 flavobacter
10	34	70.8	811	1	RD54_SCHPO	Q09772 schizosacch
11	34	70.8	989	1	T100_HUMAN	O75448 homo sapien
12	33	68.8	366	1	GHSR_PIG	Q95254 sus scrofa
13	33	68.8	439	1	PKS1_ARATH	Q9swil arabidopsis
14	33	68.8	457	1	Y162_UREPA	Q9pqy2 ureaplasma
15	33	68.8	547	1	RM56_HUMAN	P83111 homo sapien
16	33	68.8	551	1	RM56_MOUSE	Q9ep89 mus musculu
17	33	68.8	583	1	HASS_XENLA	O57428 xenopus lae

18	33	68.8	641	1	LICR_BACSU	P46321	bacillus su
19	33	68.8	702	1	DDX4_MOUSE	Q61496	mus musculu
20	33	68.8	713	1	DDX4_RAT	Q64060	rattus norv
21	33	68.8	743	1	YN25_SCHPO	O59741	schizosacch
22	32	66.7	137	1	CT85_HUMAN	Q9hlp6	homo sapien
23	32	66.7	246	1	Y350_ARATH	Q9ma63	arabidopsis
24	32	66.7	250	1	M48B_HUMAN	Q9by19	homo sapien
25	32	66.7	286	1	YDBC_ECOLI	P25906	escherichia
26	32	66.7	306	1	OPRR_PSEAE	Q01610	pseudomonas
27	32	66.7	348	1	YLW1_CAEEL	P34404	caenorhabdi
28	32	66.7	364	1	CARA_LACPL	P77885	lactobacill
29	32	66.7	375	1	CISY_MYCSM	P26491	mycobacteri
30	32	66.7	400	1	VMSA_HPBV9	P17101	hepatitis b
31	32	66.7	469	1	SYFA_SULTO	Q971d7	sulfolobus
32	32	66.7	483	1	GLYC_HUMAN	P34896	homo sapien
33	32	66.7	483	1	GLYC_RABIT	P07511	oryctolagus
34	32	66.7	483	1	GLYC_SHEEP	P35623	ovis aries
35	32	66.7	515	1	CP51_PENIT	Q12664	penicillium
36	32	66.7	548	1	NRM1_BISBI	Q95102	bison bison
37	32	66.7	558	1	EXO2_BPR69	O64299	bacterioph
38	32	66.7	700	1	HPS1_HUMAN	Q92902	homo sapien
39	32	66.7	721	1	CATB_EMENI	P78619	emericella
40	32	66.7	747	1	YFGF_ECOLI	P77172	escherichia
41	32	66.7	756	1	SEC6_HUMAN	O60645	homo sapien
42	32	66.7	1036	1	Y946_ARCFU	O29316	archaeoglob
43	32	66.7	1237	1	NME3_RAT	Q00961	rattus norv
44	31	64.6	146	1	YUAA_ECOLI	Q9jmt9	escherichia
45	31	64.6	319	1	TFB6_HALN1	Q9hsb3	halobacteri

# ALIGNMENTS

## RESULT 1

### YCF2\_MARPO

ID YCF2\_MARPO STANDARD; PRT; 2136 AA.  
AC P09975;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical 259.9 kDa protein ycf2 (ORF 2136).  
GN YCF2.  
OS Marchantia polymorpha (Liverwort).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;  
OC Marchantiopsida; Marchantiidae; Marchantiales; Marchantiineae;  
OC Marchantiaceae; Marchantia.  
OX NCBI\_TaxID=3197;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Ohyama K., Fukuzawa H., Kohchi T., Shirai H., Sano T., Sano S.,  
RA Umesono K., Shiki Y., Takeuchi M., Chang Z., Aota S., Inokuchi H.,  
RA Ozeki H.;  
RT "Chloroplast gene organization deduced from complete sequence of  
RT liverwort Marchantia polymorpha chloroplast DNA.";  
RL Nature 322:572-574(1986).  
CC -!- SIMILARITY: Belongs to the ycf2 family.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04465; CAA28078.1; -.
DR PIR; S01591; A05037.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR Pfam; PF00004; AAA; 1.
DR SMART; SM00382; AAA; 1.
KW Chloroplast; Hypothetical protein.
SQ SEQUENCE 2136 AA; 259911 MW; 5BD170C9CCF61197 CRC64;

```

```

Query Match          77.1%; Score 37; DB 1; Length 2136;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      2 ESYVPLF 8
        |||||
Db      1420 ESYVPLF 1426

```

# RESULT 2

```

YAB4_SCHPO
ID YAB4_SCHPO STANDARD; PRT; 301 AA.
AC Q09806;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C2G11.04 in chromosome I.
GN SPAC2G11.04.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

```

RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of *Schizosaccharomyces pombe*.";  
 RL Nature 415:871-880(2002).

CC -!- SIMILARITY: Contains 1 G-patch domain.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; Z54354; CAA91169.1; -.  
 DR PIR; T38570; S62459.  
 DR GeneDB\_SPombe; SPAC2G11.04; -.  
 DR InterPro; IPR000467; G\_patch.  
 DR Pfam; PF01585; G\_patch; 1.  
 DR SMART; SM00443; G\_patch; 1.  
 DR PROSITE; PS50174; G\_PATCH; 1.  
 KW Hypothetical protein.  
 FT DOMAIN 252 301 G-PATCH.  
 SQ SEQUENCE 301 AA; 33413 MW; 9754D531C339D037 CRC64;

Query Match 72.9%; Score 35; DB 1; Length 301;  
 Best Local Similarity 77.8%; Pred. No. 15;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
 ||| ||||  
 Db 153 VELYDPLFP 161

#### RESULT 3

PYRC\_SCHPO

ID PYRC\_SCHPO STANDARD; PRT; 337 AA.

AC Q9UTI0;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Probable dihydroorotase (EC 3.5.2.3) (DHOase).

GN SPAC16.03C.

OS *Schizosaccharomyces pombe* (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC *Schizosaccharomyces*.

OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of *Schizosaccharomyces pombe*.";  
 RL Nature 415:871-880(2002).  
 CC -!- CATALYTIC ACTIVITY: (S)-dihydroorotate + H(2)O = N-carbamoyl-L-  
 CC aspartate.  
 CC -!- COFACTOR: Binds 2 zinc ions per subunit (By similarity).  
 CC -!- PATHWAY: Pyrimidine biosynthesis; third step.  
 CC -!- SIMILARITY: Belongs to the DHOase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AL121745; CAB57401.1; -.  
 DR PIR; T37731; T37731.  
 DR GeneDB\_SPombe; SPAC16.03c; -.  
 DR InterPro; IPR006680; Amidohydro\_1.  
 DR InterPro; IPR004721; Pept\_M38\_DHodimr.  
 DR InterPro; IPR002195; Pept\_M38\_nph.  
 DR Pfam; PF01979; Amidohydro\_1; 1.  
 DR TIGRFAMs; TIGR00856; pyrC\_dimer; 1.  
 DR PROSITE; PS00482; DIHYDROOROTASE\_1; FALSE\_NEG.  
 DR PROSITE; PS00483; DIHYDROOROTASE\_2; 1.  
 KW Pyrimidine biosynthesis; Hydrolase; Metal-binding; Zinc.  
 FT METAL 12 12 ZINC 1 (BY SIMILARITY).

FT METAL 14 14 ZINC 1 (BY SIMILARITY).  
 FT METAL 95 95 ZINC 1 AND 2 (BY SIMILARITY).  
 FT METAL 132 132 ZINC 2 (BY SIMILARITY).  
 FT METAL 170 170 ZINC 2 (BY SIMILARITY).  
 FT METAL 240 240 ZINC 1 (BY SIMILARITY).  
 SQ SEQUENCE 337 AA; 37552 MW; 1E0D00F457A6FA02 CRC64;

Query Match 72.9%; Score 35; DB 1; Length 337;  
 Best Local Similarity 66.7%; Pred. No. 17;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VESYVPLFP 9  
 |||| | :|  
 Db 109 VESYEPFYP 117

#### RESULT 4

SYI\_CHLTR

ID SYI\_CHLTR STANDARD; PRT; 1036 AA.  
 AC 084022;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Isoleucyl-tRNA synthetase (EC 6.1.1.5) (Isoleucine--tRNA ligase)  
 DE (IleRS).  
 GN ILES OR CT019.  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D/UW-3/Cx;  
 RX MEDLINE=99000809; PubMed=9784136;  
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
 RA Davis R.W.;  
 RT "Genome sequence of an obligate intracellular pathogen of humans:  
 RT Chlamydia trachomatis.";  
 RL Science 282:754-759(1998).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-isoleucine + tRNA(Ile) = AMP +  
 CC diphosphate + L-isoleucyl-tRNA(Ile).  
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -!- SUBUNIT: Monomer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE001277; AAC67609.1; -.  
 DR PIR; F71565; F71565.  
 DR HSSP; P56690; 1ILE.

DR InterPro; IPR002300; tRNA-synt\_1a.  
 DR InterPro; IPR001412; tRNA-synt\_I.  
 DR InterPro; IPR002301; tRNA-synt\_ile.  
 DR InterPro; IPR009008; ValRS\_IleRS\_edit.  
 DR Pfam; PF00133; tRNA-synt\_1; 1.  
 DR PRINTS; PR00984; TRNASYNTHILE.  
 DR TIGRFAMs; TIGR00392; ileS; 1.  
 DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_I; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Metal-binding; Zinc; Complete proteome.  
 FT SITE 46 56 "HIGH" REGION.  
 FT SITE 589 593 "KMSKS" REGION.  
 FT BINDING 592 592 ATP (BY SIMILARITY).  
 SQ SEQUENCE 1036 AA; 118985 MW; DC4420FD37343EFD CRC64;

Query Match 72.9%; Score 35; DB 1; Length 1036;  
 Best Local Similarity 75.0%; Pred. No. 58;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9  
 :|| ||||  
 Db 284 QSYEPLFP 291

#### RESULT 5

SA2\_XENLA

ID SA2\_XENLA STANDARD; PRT; 1194 AA.

AC Q9DGN0;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Cohesin subunit SA-2 (XSA-2) (Stromal antigen 2 homolog) (SCC3 homolog 2).

GN STAG2 OR XSA2.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI\_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A., AND IDENTIFICATION IN A COHESIN COMPLEX WITH  
 RP SMC1; SMC3 AND RAD21.

RC TISSUE=Egg;

RX MEDLINE=20391966; PubMed=10931856;

RA Losada A., Yokochi T., Kobayashi R., Hirano T.;

RT "Identification and characterization of SA/Scp3p subunits in the  
 RT Xenopus and human cohesin complexes.";

RL J. Cell Biol. 150:405-416(2000).

RN [2]

RP PHOSPHORYLATION BY PLK.

RX MEDLINE=21930342; PubMed=11931760;

RA Sumara I., Vorlaufer E., Stukenberg P.T., Kelm O., Redemann N.,

RA Nigg E.A., Peters J.-M.;

RT "The dissociation of cohesin from chromosomes in prophase is regulated  
 RT by Polo-like kinase.";

RL Mol. Cell 9:515-525(2002).

CC -!- FUNCTION: Component of cohesin complex, a complex required for the



cohesion of sister chromatids after DNA replication. The cohesin complex apparently forms a large proteinaceous ring within which sister chromatids can be trapped. At anaphase, the complex is cleaved and dissociates from chromatin, allowing sister chromatids to segregate. The cohesin complex may also play a role in spindle pole assembly during mitosis.

CC -!- SUBUNIT: Interacts directly with RAD21 in cohesin complex. Cohesin complexes are composed of a heterodimer between SMC1 and SMC3, which are attached via their hinge domain, and RAD21 which link them at their heads, and one STAG protein (STAG1 or STAG2). In cohesin complexes, STAG1 is mutually exclusive with STAG2.

CC -!- SUBCELLULAR LOCATION: Nuclear protein. Associates with chromatin. Before prophase it is scattered along chromosome arms. During prophase, most of cohesin complexes dissociate from chromatin probably because of phosphorylation by PLK, except at centromeres, where cohesin complexes remain. At anaphase, the RAD21 subunit of cohesin is cleaved, leading to the dissociation of the complex from chromosomes, allowing chromosome separation.

CC -!- PTM: Phosphorylated by PLK. The large dissociation of cohesin from chromosome arms during prophase is partly due to its phosphorylation.

CC -!- SIMILARITY: Belongs to the SCC3 family.

CC -!- SIMILARITY: Contains 1 SCD (stromalin conservative) domain.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AF255018; AAG00431.1; -.

DR InterPro; IPR008938; ARM.

DR InterPro; IPR005032; STAG.

DR Pfam; PF03365; STAG; 1.

KW Mitosis; Cell cycle; Chromosome partition; Nuclear protein;

KW Phosphorylation.

FT DOMAIN 224 309 SCD.

FT DOMAIN 606 609 POLY-GLU.

FT DOMAIN 1021 1026 POLY-SER.

SQ SEQUENCE 1194 AA; 137824 MW; 804FC73E468B13CE CRC64;

Query Match 72.9%; Score 35; DB 1; Length 1194;

Best Local Similarity 62.5%; Pred. No. 68;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9

|||:|::|

Db 1061 ESYMPVYP 1068

RESULT 6

SRPP\_HEVBR

ID SRPP\_HEVBR STANDARD; PRT; 204 AA.

AC O82803; Q9S8I3; Q9S8I4; Q9S8I5; Q9S9G7;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Small rubber particle protein (SRPP) (22 kDa rubber particle protein)  
 DE (22 kDa RPP) (Latex allergen Hev b 3) (27 kDa natural rubber  
 DE allergen).  
 GN SRPP OR HEVB3.  
 OS Hevea brasiliensis (Para rubber tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Malpighiales; Euphorbiaceae; Crotonoideae; Micrandreae;  
 OC Hevea.  
 OX NCBI\_TaxID=3981;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. RRIM 600; TISSUE=Latex;  
 RX MEDLINE=99287916; PubMed=10358068;  
 RA Oh S.K., Kang H., Shin D.H., Yang J., Chow K.-S., Yeang H.Y.,  
 RA Wagner B., Breiteneder H., Han K.-H.;  
 RT "Isolation, characterization, and functional analysis of a novel cDNA  
 RT clone encoding a small rubber particle protein from Hevea  
 RT brasiliensis.";  
 RL J. Biol. Chem. 274:17132-17138(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. RRIM 600; TISSUE=Latex;  
 RX MEDLINE=20020199; PubMed=10550757;  
 RA Wagner B., Krebitz M., Buck D., Niggemann B., Yeang H.Y., Han K.-H.,  
 RA Scheiner O., Breiteneder H.;  
 RT "Cloning, expression, and characterization of recombinant Hev b 3, a  
 RT Hevea brasiliensis protein associated with latex allergy in patients  
 RT with spina bifida.";  
 RL J. Allergy Clin. Immunol. 104:1084-1092(1999).  
 RN [3]  
 RP SEQUENCE OF 12-17; 23-31; 36-50; 54-68; 73-86; 91-104; 106-123;  
 RP 125-143 AND 175-192.  
 RC TISSUE=Latex;  
 RX MEDLINE=95195373; PubMed=7888787;  
 RA Alenius H., Kalkkinen N., Lukka M., Turjanmaa K., Reunala T.,  
 RA Maekinen-Kiljunen S., Palosuo T.;  
 RT "Purification and partial amino acid sequencing of a 27-kD natural  
 RT rubber allergen recognized by latex-allergic children with spina  
 RT bifida.";  
 RL Int. Arch. Allergy Immunol. 106:258-262(1995).  
 RN [4]  
 RP SEQUENCE OF 8-33; 36-49; 54-68; 73-90 AND 178-194.  
 RC TISSUE=Latex;  
 RX MEDLINE=95378697; PubMed=7650400;  
 RA Lu L.-J., Kurup V.P., Hoffman D.R., Kelly K.J., Murali P.S.,  
 RA Fink J.N.;  
 RT "Characterization of a major latex allergen associated with  
 RT hypersensitivity in spina bifida patients.";  
 RL J. Immunol. 155:2721-2728(1995).  
 RN [5]  
 RP SEQUENCE OF 18-35; 54-68; 73-90; 106-122 AND 183-194, AND MASS  
 RP SPECTROMETRY.  
 RC STRAIN=cv. RRIM 600, and cv. Costa Rica; TISSUE=Latex;  
 RX MEDLINE=98300440; PubMed=9636811;

RA Yeang H.Y., Ward M.A., Zamri A.S.M., Dennis M.S., Light D.R.;  
 RT "Amino acid sequence similarity of Hev b 3 to two previously reported  
 RT 27- and 23-kDa latex proteins allergenic to spina bifida patients.";  
 RL Allergy 53:513-519(1998).  
 RN [6]  
 RP CHARACTERIZATION OF EPITOPES.  
 RX MEDLINE=21173202; PubMed=11275264;  
 RA Banerjee B., Kanitpong K., Fink J.N., Zussman M., Sussman G.L.,  
 RA Kelly K.J., Kurup V.P.;  
 RT "Unique and shared IgE epitopes of Hev b 1 and Hev b 3 in latex  
 RT allergy.";  
 RL Mol. Immunol. 37:789-798(2000).  
 CC -!- FUNCTION: Involved in the biosynthesis of rubber, an isoprenoid  
 CC polymer (cis-1,4-polyisoprene).  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC. FOUND IN LATEX, THE CYTOPLASM  
 CC OF LATICIFER CELLS. TIGHTLY BOUND TO SMALL RUBBER PARTICLES.  
 CC -!- TISSUE SPECIFICITY: Highly expressed in the specialized vessel  
 CC laticifers. Also detected in leaves.  
 CC -!- INDUCTION: Not induced by wounding or etephon treatment.  
 CC -!- PTM: The N-terminus is blocked.  
 CC -!- PTM: THE CONSISTENT GAPS OF ABOUT 266 DA OBSERVED BY MS IN VARIOUS  
 CC FORMS OF THE INTACT PROTEIN SUGGEST THE ADDITION OF STEAROLYL  
 CC GROUPS.  
 CC -!- MASS SPECTROMETRY: MW=22258; METHOD=Electrospray.  
 CC -!- MASS SPECTROMETRY: MW=22524; METHOD=Electrospray.  
 CC -!- MASS SPECTROMETRY: MW=22791; METHOD=Electrospray.  
 CC -!- MASS SPECTROMETRY: MW=23058; METHOD=Electrospray.  
 CC -!- ALLERGEN: Causes an allergic reaction in human. Involved in latex  
 CC allergic reactions. Major spina bifida (SB) associated latex  
 CC allergen. Patients with SB are especially sensitive to specific  
 CC allergens that do not usually affect adult patients without SB who  
 CC are otherwise sensitive to latex. These patients are multioperated  
 CC patients, whose mucosal membranes have been frequently in contact  
 CC with latex products. Many patients with SB are sensitized to low  
 CC molecular weight polypeptides derived from the fragmentation of  
 CC Hev b 3.  
 CC -!- SIMILARITY: Belongs to the REF/SRPP family.  
 CC -!- DATABASE: NAME=Protein Spotlight;  
 CC NOTE=Issue 13 of August 2001;  
 CC WWW="http://www.expasy.org/spotlight/articles/sptlt013.html".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF051317; AAC82355.1; -.  
 DR EMBL; AJ223388; CAA11303.1; -.  
 DR EMBL; AJ223389; CAA11304.1; -.  
 DR EMBL; AJ223390; CAA11305.1; -.  
 DR PIR; T10766; T10766.  
 DR InterPro; IPR008802; REF.  
 DR Pfam; PF05755; REF; 1.  
 KW Allergen.

FT CONFLICT 39 39 G -> Y (IN REF. 4).  
 FT CONFLICT 44 44 G -> P (IN REF. 4).  
 FT CONFLICT 49 49 E -> K (IN REF. 4).  
 FT CONFLICT 78 78 V -> T OR P (IN REF. 5; CV. COSTA RICA).  
 FT CONFLICT 118 122 GVQEG -> RKEKK (IN REF. 5; CV. RRIMM  
 FT 600).  
 FT CONFLICT 192 192 T -> G (IN REF. 3).  
 SQ SEQUENCE 204 AA; 22344 MW; 2A3BAE7D8325D833 CRC64;

Query Match 70.8%; Score 34; DB 1; Length 204;  
 Best Local Similarity 66.7%; Pred. No. 16;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
 | ||:| |  
 Db 183 VSSYLPLLP 191

# RESULT 7

## OTC\_PYRAE

ID OTC\_PYRAE STANDARD; PRT; 304 AA.  
 AC Q8ZW40;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ornithine carbamoyltransferase (EC 2.1.3.3) (OTCase).  
 GN ARGF OR PAE1985.  
 OS Pyrobaculum aerophilum.  
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;  
 OC Thermoproteaceae; Pyrobaculum.  
 OX NCBI\_TaxID=13773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;  
 RX PubMed=11792869;  
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,  
 RA Miller J.H.;  
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum  
 RT aerophilum."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).  
 CC -!- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate  
 CC + L-citrulline.  
 CC -!- PATHWAY: Arginine biosynthesis; sixth step.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -!- SIMILARITY: Belongs to the ATCase/OTCase family.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AE009850; AAL63862.1; -.  
 DR HAMAP; MF\_01109; -; 1.  
 DR InterPro; IPR006130; Asp/Orn\_COtranf.

DR InterPro; IPR002292; Orn\_carbtransf.  
 DR InterPro; IPR006131; OTCace\_O.  
 DR InterPro; IPR006132; OTCace\_P.  
 DR Pfam; PF00185; OTCace; 1.  
 DR Pfam; PF02729; OTCace\_N; 1.  
 DR PRINTS; PR00100; AOTCASE.  
 DR TIGRFAMs; TIGR00658; orni\_carb\_tr; 1.  
 DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.  
 KW Arginine biosynthesis; Transferase; Complete proteome.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT SITE 25 25 IMPORTANT FOR STRUCTURAL INTEGRITY (BY  
 FT SIMILARITY).  
 FT SITE 51 55 CARBAMOYLPHOSPHATE BINDING (BY  
 FT SIMILARITY).  
 FT SITE 102 102 CARBAMOYLPHOSPHATE BINDING (BY  
 FT SIMILARITY).  
 FT SITE 129 129 CARBAMOYLPHOSPHATE BINDING (BY  
 FT SIMILARITY).  
 FT SITE 142 142 IMPORTANT FOR STRUCTURAL INTEGRITY (BY  
 FT SIMILARITY).  
 FT SITE 262 265 ORNITHINE BINDING (BY SIMILARITY).  
 SQ SEQUENCE 304 AA; 34110 MW; DA73031506362E9B CRC64;

Query Match 70.8%; Score 34; DB 1; Length 304;  
 Best Local Similarity 75.0%; Pred. No. 25;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9  
 | | | | |  
 Db 32 EIYTPLFP 39

# RESULT 8

## AMIC\_PSEAE

ID AMIC\_PSEAE STANDARD; PRT; 384 AA.  
 AC P27017;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Aliphatic amidase expression-regulating protein.  
 GN AMIC OR PA3364.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-18.  
 RC STRAIN=PAC;  
 RX MEDLINE=91317707; PubMed=1907262;  
 RA Wilson S.A., Drew R.E.;  
 RT "Cloning and DNA sequence of *amiC*, a new gene regulating expression  
 RT of the *Pseudomonas aeruginosa* aliphatic amidase, and purification of  
 RT the *amiC* product."  
 RL J. Bacteriol. 173:4914-4921(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PA01;

RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an  
 RT opportunistic pathogen.";  
 RL Nature 406:959-964(2000).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY.  
 RX MEDLINE=92106343; PubMed=1762155;  
 RA Wilson S.A., Chayen N.E., Hemmings A.M., Drew R.E., Pearl L.H.;  
 RT "Crystallization of and preliminary X-ray data for the negative  
 RT regulator (AmiC) of the amidase operon of *Pseudomonas aeruginosa*.";  
 RL J. Mol. Biol. 222:869-871(1991).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS), AND REVISIONS TO 26-27.  
 RX MEDLINE=95112789; PubMed=7813419;  
 RA Pearl L.H., O'Hara B.P., Drew R.E., Wilson S.A.;  
 RT "Crystal structure of AmiC: the controller of transcription  
 RT antitermination in the amidase operon of *Pseudomonas aeruginosa*.";  
 RL EMBO J. 13:5810-5817(1994).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF COMPLEX WITH AMIR.  
 RC STRAIN=PAC1;  
 RX MEDLINE=99437995; PubMed=10508151;  
 RA O'Hara B.P., Norman R.A., Wan P.T., Roe S.M., Barrett T.E., Drew R.E.,  
 RA Pearl L.H.;  
 RT "Crystal structure and induction mechanism of AmiC-AmiR: a  
 RT ligand-regulated transcription antitermination complex.";  
 RL EMBO J. 18:5175-5186(1999).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).  
 RC STRAIN=PAC1;  
 RX MEDLINE=20175740; PubMed=10708652;  
 RA O'Hara B.P., Wilson S.A., Lee A.W., Roe S.M., Siligardi G., Drew R.E.,  
 RA Pearl L.H.;  
 RT "Structural adaptation to selective pressure for altered ligand  
 RT specificity in the *Pseudomonas aeruginosa* amide receptor, AmiC.";  
 RL Protein Eng. 13:129-132(2000).  
 CC -!- FUNCTION: NEGATIVELY REGULATES THE EXPRESSION OF THE ALIPHATIC  
 CC AMIDASE OPERON. AMIC FUNCTIONS BY INHIBITING THE ACTION OF AMIR AT  
 CC THE PROTEIN LEVEL. IT EXHIBITS PROTEIN KINASE ACTIVITY.  
 CC -!- SUBUNIT: Homodimer. Forms a complex with amiR.  
 CC -!- DOMAIN: CONSISTS OF TWO BETA-ALPHA-BETA DOMAINS WITH A CENTRAL  
 CC CLEFT IN WHICH THE AMIDE BINDS.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X13776; CAA32024.1; -.  
 DR EMBL; AE004758; AAG06752.1; -.  
 DR PIR; C83226; C83226.  
 DR PDB; 1PEA; 03-APR-96.  
 DR PDB; 1QNL; 23-DEC-99.  
 DR PDB; 1QO0; 23-DEC-99.  
 DR InterPro; IPR000709; Leu\_Ile\_Val\_bind.  
 DR PRINTS; PR00337; LEUILEVALBP.  
 KW Transferase; Kinase; Repressor; 3D-structure; Complete proteome.  
 FT INIT\_MET 0 0  
 FT VARIANT 105 105 T -> N (IN STRAIN PAC181; BUTYRAMIDE  
 FT INDUCIBLE PHENOTYPE).  
 FT CONFLICT 26 27 QR -> HA (IN REF. 1).  
 FT CONFLICT 185 185 V -> L (IN REF. 1).  
 FT CONFLICT 262 262 A -> P (IN REF. 1).  
 FT CONFLICT 304 304 S -> N (IN REF. 1).  
 FT CONFLICT 318 318 C -> D (IN REF. 1).  
 FT CONFLICT 382 382 A -> P (IN REF. 1).  
 FT STRAND 8 12  
 FT TURN 18 19  
 FT HELIX 20 38  
 FT TURN 39 42  
 FT STRAND 43 43  
 FT TURN 44 45  
 FT STRAND 46 46  
 FT STRAND 49 53  
 FT TURN 56 57  
 FT HELIX 59 71  
 FT TURN 72 73  
 FT STRAND 77 80  
 FT HELIX 84 96  
 FT TURN 97 98  
 FT STRAND 100 103  
 FT TURN 114 115  
 FT STRAND 116 118  
 FT HELIX 123 125  
 FT TURN 126 126  
 FT HELIX 127 135  
 FT TURN 136 138  
 FT STRAND 141 147  
 FT HELIX 150 165  
 FT TURN 166 167  
 FT STRAND 169 176  
 FT HELIX 182 195  
 FT TURN 196 196  
 FT STRAND 199 203  
 FT TURN 206 206  
 FT HELIX 207 220  
 FT STRAND 228 230  
 FT HELIX 235 238  
 FT TURN 239 240  
 FT HELIX 243 246  
 FT TURN 247 248  
 FT STRAND 250 253  
 FT TURN 257 258  
 FT HELIX 262 272  
 FT TURN 273 274

FT	TURN	277	278
FT	HELIX	283	302
FT	TURN	303	303
FT	HELIX	307	314
FT	TURN	315	316
FT	STRAND	319	321
FT	TURN	322	323
FT	STRAND	324	328
FT	TURN	330	332
FT	STRAND	335	337
FT	STRAND	340	344
FT	TURN	346	347
FT	STRAND	350	355
FT	STRAND	360	360
FT	HELIX	368	370
SQ	SEQUENCE	384 AA;	42676 MW; 103EDFA52E406AF1 CRC64;

Query Match 70.8%; Score 34; DB 1; Length 384;  
 Best Local Similarity 62.5%; Pred. No. 32;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ESYVPLFP 9  
 | | : | | : |  
 Db 173 EIYIPLYP 180

# RESULT 9

## EFTU\_FLAPE

ID EFTU\_FLAPE STANDARD; PRT; 395 AA.  
 AC P42476;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Elongation factor Tu (EF-Tu).  
 GN TUF.  
 OS Flavobacterium ferrugineum.  
 OC Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales;  
 OC Flavobacteriaceae; Flavobacterium.  
 OX NCBI\_TaxID=249;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 13524;  
 RX MEDLINE=94368062; PubMed=8085791;  
 RA Ludwig W., Neumaier J., Klugbauer N., Brockmann E., Roller C.,  
 RA Klugbauer S., Reetz K., Schachtner I., Ludvigsen A.,  
 RA Bachleitner M., Fischer U., Schleifer K.H.;  
 RT "Phylogenetic relationships of Bacteria based on comparative sequence  
 RT analysis of elongation factor Tu and ATP-synthase beta-subunit  
 RT genes.";  
 RL Antonie Van Leeuwenhoek 64:285-305(1993).  
 CC -!- FUNCTION: This protein promotes the GTP-dependent binding of  
 CC aminoacyl-tRNA to the A-site of ribosomes during protein  
 CC biosynthesis.  
 CC -!- SUBUNIT: Monomer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.  
 CC EF-Tu/EF-1A subfamily.



```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X76867; CAA54195.1; -.
DR HSSP; P02990; 1EFU.
DR HAMAP; MF_00118; -; 1.
DR InterPro; IPR004541; EF-Tu.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR004160; EFTU_Cterm.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR009001; Elong_init_C.
DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRFAMs; TIGR00485; EF-Tu; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFACITOR_GTP; 1.
KW Elongation factor; Protein biosynthesis; GTP-binding.
FT NP_BIND      19      26      GTP (BY SIMILARITY).
FT NP_BIND      81      85      GTP (BY SIMILARITY).
FT NP_BIND     136     139      GTP (BY SIMILARITY).
SQ SEQUENCE    395 AA;  43166 MW;  0DA5F131B61DDC68 CRC64;

```

```

Query Match          70.8%; Score 34; DB 1; Length 395;
Best Local Similarity 66.7%; Pred. No. 33;
Matches      6; Conservative      2; Mismatches      1; Indels      0; Gaps      0;

```

```

QY      1 VESYVPLFP 9
        |:|:| |
Db      196 VDSYIPLPP 204

```

# RESULT 10

RD54\_SCHPO

```

ID   RD54_SCHPO      STANDARD;          PRT;   811 AA.
AC   Q09772; Q7ZA42;
DT   01-NOV-1995 (Rel. 32, Created)
DT   15-MAR-2004 (Rel. 43, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Meiotic recombination protein rdh54 (RAD54 protein homolog 2).
GN   RDH54 OR SPAC22F3.03C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A., FUNCTION, AND INTERACTION WITH RHP51 AND DMC1.

```

RX MEDLINE=22954574; PubMed=14551247;  
 RA Catlett M.G., Forsburg S.L.;  
 RT "S. pombe Rdh54 (TID1) acts with Rhp54 (RAD54) to repair meiotic  
 RT double strand breaks.";  
 RL Mol. Biol. Cell 14:0-0(2003).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe.";  
 RL Nature 415:871-880(2002).  
 CC -!- FUNCTION: Acts with rhp54 to repair meiotic double strand breaks  
 CC via homologous recombination. Involved in meiotic DNA  
 CC recombination.  
 CC -!- SUBUNIT: Interacts with rhp51 and dmcl.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- SIMILARITY: Belongs to the SNF2/RAD54 helicase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AY293737; AAP44116.1; -.  
 DR EMBL; Z54285; CAA91068.3; -.  
 DR PIR; T38188; T38188.  
 DR GeneDB\_SPombe; SPAC22F3.03c; -.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR InterPro; IPR000330; SNF2\_N.

DR Pfam; PF00271; helicase\_C; 1.  
 DR Pfam; PF00176; SNF2\_N; 1.  
 DR SMART; SM00487; DEXDc; 1.  
 DR SMART; SM00490; HELICc; 1.  
 KW DNA recombination; DNA repair; Helicase; DNA-binding; ATP-binding;  
 KW Nuclear protein.  
 FT NP\_BIND 235 242 ATP (POTENTIAL).  
 SQ SEQUENCE 811 AA; 92965 MW; C5C88292A8E93604 CRC64;

Query Match 70.8%; Score 34; DB 1; Length 811;  
 Best Local Similarity 66.7%; Pred. No. 71;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
 |||:|||  
 Db 791 VEDYIPLEP 799

# RESULT 11

## T100\_HUMAN

ID T100\_HUMAN STANDARD; PRT; 989 AA.  
 AC O75448; Q14143; Q9NNY5;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Thyroid hormone receptor-associated protein complex 100 kDa component  
 DE (Trap100) (Vitamin D3 receptor-interacting protein complex component  
 DE DRIP100) (Activator-recruited cofactor 100 kDa component) (ARC100).  
 GN TRAP100 OR DRIP100 OR ARC100 OR KIAA0130.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.  
 RX MEDLINE=98318590; PubMed=9653119;  
 RA Yuan C.-X., Ito M., Fondell J.D., Fu Z.-Y., Roeder R.G.;  
 RT "The TRAP220 component of a thyroid hormone receptor-associated  
 RT protein (TRAP) coactivator complex interacts directly with nuclear  
 RT receptors in a ligand-dependent fashion."  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:7939-7944(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Bone marrow;  
 RX MEDLINE=96127530; PubMed=8590280;  
 RA Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;  
 RT "Prediction of the coding sequences of unidentified human genes. IV.  
 RT The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by  
 RT analysis of cDNA clones from human cell line KG-1."  
 RL DNA Res. 2:167-174(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND IDENTIFICATION IN DRIP COMPLEX.  
 RX MEDLINE=98301435; PubMed=9637681;  
 RA Rachez C., Suldan Z., Ward J., Chang C.-P.B., Burakov D.,  
 RA Erdjument-Bromage H., Tempst P., Freedman L.P.;  
 RT "A novel protein complex that interacts with the vitamin D3 receptor  
 RT in a ligand-dependent manner and enhances VDR transactivation in a

RT cell-free system.";  
 RL Genes Dev. 12:1787-1800(1998).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Eye;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [5]  
 RP IDENTIFICATION IN ARC COMPLEX, AND SEQUENCE OF 1-11 AND 957-965.  
 RX MEDLINE=99249346; PubMed=10235267;  
 RA Naeae A.M., Beaurang P.A., Zhou S., Abraham S., Solomon W.B.,  
 RA Tjian R.;  
 RT "Composite co-activator ARC mediates chromatin-directed  
 RT transcriptional activation.";  
 RL Nature 398:828-832(1999).  
 CC -!- FUNCTION: Plays a role in transcriptional coactivation. Plays a  
 CC role in nuclear hormone receptor-mediated transactivation.  
 CC -!- SUBUNIT: Subunit of the large multiprotein complexes ARC/DRIP and  
 CC TRAP. Interacts and co-precipitates with Trap220. Does not  
 CC directly bind nuclear hormone receptors.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- TISSUE SPECIFICITY: Ubiquitous. Abundant in skeletal muscle, heart  
 CC and placenta.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF055995; AAC39855.1; -.  
 DR EMBL; D50920; BAA09479.1; -.  
 DR EMBL; AF277379; AAF78764.1; -.  
 DR EMBL; BC011375; AAH11375.1; -.  
 DR MIM; 607000; -.  
 DR GO; GO:0000119; C:mediator complex; NAS.

DR GO; GO:0005634; C:nucleus; IDA.  
 DR GO; GO:0030374; F:ligand-dependent nuclear receptor transcrip. . .; NAS.  
 DR GO; GO:0004872; F:receptor activity; IDA.  
 DR GO; GO:0016455; F:RNA polymerase II transcription mediator ac. . .; IDA.  
 DR GO; GO:0046966; F:thyroid hormone receptor binding; IDA.  
 DR GO; GO:0003712; F:transcription cofactor activity; IDA.  
 DR GO; GO:0016563; F:transcriptional activator activity; IDA.  
 DR GO; GO:0042809; F:vitamin D receptor binding; NAS.  
 DR GO; GO:0030521; P:androgen receptor signaling pathway; IDA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.  
 DR GO; GO:0006367; P:transcription initiation from Pol II promoter; IDA.  
 KW Transcription regulation; Zinc-finger; Repeat; ATP-binding;  
 KW Nuclear protein.  
 FT DOMAIN 128 861 6 X REPEATS OF L-X-X-L-L.  
 FT REPEAT 128 132 1.  
 FT REPEAT 344 348 2.  
 FT REPEAT 448 452 3.  
 FT REPEAT 557 561 4.  
 FT REPEAT 788 792 5.  
 FT REPEAT 857 861 6.  
 FT ZN\_FING 93 117 POTENTIAL.  
 FT NP\_BIND 440 447 ATP (POTENTIAL).  
 FT CONFLICT 20 20 D -> Y (IN REF. 2).  
 FT CONFLICT 204 204 A -> T (IN REF. 2).  
 FT CONFLICT 555 555 E -> G (IN REF. 3).  
 SQ SEQUENCE 989 AA; 110304 MW; CCEDE7D4E74D890C CRC64;

Query Match 70.8%; Score 34; DB 1; Length 989;  
 Best Local Similarity 55.6%; Pred. No. 88;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
 :| |: |||  
 Db 840 IEDYISLFP 848

# RESULT 12

## GHSR\_PIG

ID GHSR\_PIG STANDARD; PRT; 366 AA.  
 AC Q95254; Q95255;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing peptide receptor) (GHRP) (Ghrelin receptor).  
 GN GHSR.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).  
 RC STRAIN=Yorkshire; TISSUE=Pituitary;  
 RX MEDLINE=96337998; PubMed=8688086;  
 RA Howard A.D., Feighner S.D., Cully D.F., Arena J.P.,  
 RA Liberator P.A., Rosenblum C.I., Hamelin M., Hreniuk D.L.,  
 RA Palyha O.C., Anderson J., Paress P.S., Diaz C., Chou M., Liu K.K.,

RA McKee K.K., Pong S.-S., Chaung L.-Y., Elbrecht A., Dashkevicz M.,  
 RA Heavens R., Rigby M., Sirinathsinghji D.J.S., Dean D.C., Melillo D.G.,  
 RA Patchett A.A., Nargund R., Griffin P.R., Demartino J.A., Gupta S.K.,  
 RA Schaeffer J.M., Smith R.G., van der Ploeg L.H.T.;  
 RT "A receptor in pituitary and hypothalamus that functions in growth  
 RT hormone release.";  
 RL Science 273:974-977(1996).  
 CC -!- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.  
 CC Stimulates growth hormone secretion. Binds also other growth  
 CC hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6)  
 CC as well as non-peptide, low molecular weight secretagogues (e.g.  
 CC L-692,429, MK-0677, adenosine).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1A;  
 CC IsoId=Q95254-1; Sequence=Displayed;  
 CC Name=1B;  
 CC IsoId=Q95254-2; Sequence=VSP\_001918, VSP\_001919;  
 CC -!- TISSUE SPECIFICITY: Pituitary and hypothalamus.  
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U60178; AAC48630.1; -.  
 DR EMBL; U60180; AAC48631.1; -.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Alternative splicing.  
 FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 41 66 1 (POTENTIAL).  
 FT DOMAIN 67 72 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 73 96 2 (POTENTIAL).  
 FT DOMAIN 97 117 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 118 139 3 (POTENTIAL).  
 FT DOMAIN 140 162 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 163 183 4 (POTENTIAL).  
 FT DOMAIN 184 211 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 212 235 5 (POTENTIAL).  
 FT DOMAIN 236 263 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 264 285 6 (POTENTIAL).  
 FT DOMAIN 286 302 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 303 326 7 (POTENTIAL).  
 FT DOMAIN 327 366 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 116 198 BY SIMILARITY.  
 FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT VARSPLIC 266 289 AVVVFAFILCWLPFHVGRYLFSSKS -> GGSQCALELSLPG  
 FT PLHSSCLFSSP (in isoform 1B).  
 FT /FTId=VSP\_001918.  
 FT VARSPLIC 290 366 Missing (in isoform 1B).  
 FT /FTId=VSP\_001919.  
 SQ SEQUENCE 366 AA; 41194 MW; 2C850B3EF61B7C1C CRC64;

Query Match 68.8%; Score 33; DB 1; Length 366;  
 Best Local Similarity 66.7%; Pred. No. 48;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
 || :||||  
 Db 31 VEELLPLFP 39

# RESULT 13

## PKS1\_ARATH

ID PKS1\_ARATH STANDARD; PRT; 439 AA.  
 AC Q9SWI1; O80610;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Phytochrome kinase substrate 1.  
 GN PKS1 OR AT2G02950 OR T17M13.12.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=99278532; PubMed=10348744;  
 RA Fankhauser C., Yeh K.C., Lagarias J.C., Zhang H., Elich T.D.,  
 RA Chory J.;  
 RT "PKS1, a substrate phosphorylated by phytochrome that modulates light  
 RT signaling in Arabidopsis.";  
 RL Science 284:1539-1541(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=20083487; PubMed=10617197;  
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,  
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,  
 RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,  
 RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,  
 RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,  
 RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,  
 RA Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 402:761-768(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;

RX MEDLINE=22954850; PubMed=14593172;  
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,  
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,  
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,  
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,  
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,  
 RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,  
 RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,  
 RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,  
 RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,  
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RT "Empirical analysis of transcriptional activity in the Arabidopsis  
 RT genome.";  
 RL Science 302:842-846(2003).  
 CC -!- FUNCTION: May be responsible for light-regulated cytoplasmic  
 CC sequestration of phytochromes or may be a negative regulator of  
 CC phytochrome B signaling.  
 CC -!- SUBUNIT: Interacts with the C-termini of both phytochromes A  
 CC (phyA) and B (phyB). Binds both spectral forms of phytochrome, Pr  
 CC and Pfr.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- TISSUE SPECIFICITY: Expressed in young seedlings in both darkness  
 CC and light. Moderate in leaves and very low in roots and flowers.  
 CC -!- DEVELOPMENTAL STAGE: Decreases with development.  
 CC -!- PTM: Phosphorylated on Ser and to a lower extent on Thr by  
 CC phytochromes. Phosphorylation is stimulated twofold by red light.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF149053; AAD38033.1; -.  
 DR EMBL; AC004138; AAC32913.1; -.  
 DR EMBL; AF325064; AAK17132.1; -.  
 DR EMBL; AY052708; AAK96612.1; -.  
 DR EMBL; AY063721; AAL36071.1; -.  
 DR PIR; E84442; E84442.  
 DR PIR; T52304; T52304.  
 KW Phosphorylation; Phytochrome signaling pathway.  
 FT DOMAIN 31 34 POLY-SER.  
 FT DOMAIN 365 369 POLY-SER.  
 FT CONFLICT 196 196 E -> D (IN REF. 1).  
 SQ SEQUENCE 439 AA; 48084 MW; 7A0A825FF7D3F3FC CRC64;

Query Match 68.8%; Score 33; DB 1; Length 439;  
 Best Local Similarity 85.7%; Pred. No. 58;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 SYVPLFP 9  
 |||| ||  
 Db 402 SYVPRFP 408



RESULT 14

Y162\_UREPA

ID Y162\_UREPA STANDARD; PRT; 457 AA.

AC Q9PQY2;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein UU162.

GN UU162.

OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.

OX NCBI\_TaxID=134821;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Serovar 3;

RX MEDLINE=20500219; PubMed=11048724;

RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,

RA Cassell G.H.;

RT "The complete sequence of the mucosal pathogen Ureaplasma

RT urealyticum.";

RL Nature 407:757-762(2000).

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

DR EMBL; AE002116; AAF30568.1; -.

DR InterPro; IPR008160; Collagen.

DR Pfam; PF01391; Collagen; 1.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 457 AA; 51904 MW; 57D0EFC8D0A4D637 CRC64;

Query Match 68.8%; Score 33; DB 1; Length 457;

Best Local Similarity 66.7%; Pred. No. 61;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9

|| |: |||

Db 9 VEEYLSLFP 17

RESULT 15

RM56\_HUMAN

ID RM56\_HUMAN STANDARD; PRT; 547 AA.

AC P83111; P83096;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Mitochondrial 39S ribosomal protein L56 (MRP-L56) (Serine beta

DE lactamase-like protein LACTB).

GN LACTB.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RA Li W.B., Gruber C., Jessee J., Polayes D.;  
 RT "Full-length cDNA libraries and normalization.";  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Birren B., Linton L., Nusbaum C., Lander E.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE OF 1-373 FROM N.A.  
 RC TISSUE=Placenta;  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
 RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
 RA Ninomiya K., Iwayanagi T.;  
 RT "NEDO human cDNA sequencing project.";  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [4]  
 RP IDENTIFICATION, AND CONCEPTUAL TRANSLATION.  
 RA Koc E.C., Burkhart W., Blackburn K., Schlatzer D.M., Moseley A.,  
 RA Spremulli L.L.;  
 RT "The large subunit of the mammalian mitochondrial ribosome:  
 RT Identification of the full complement of ribosomal proteins present.";  
 RL Submitted (AUG-2001) to Swiss-Prot.  
 RN [5]  
 RP IDENTIFICATION, AND CONCEPTUAL TRANSLATION.  
 RX MEDLINE=21564197; PubMed=11707067;  
 RA Smith T.S., Southan C., Ellington K., Campbell D., Tew D.G.,  
 RA Debouck C.;  
 RT "Identification, genomic organization, and mRNA expression of LACTB,  
 RT encoding a serine beta-lactamase-like protein with an amino-terminal  
 RT transmembrane domain.";  
 RL Genomics 78:12-14(2001).  
 CC -!- SUBCELLULAR LOCATION: Mitochondrial (By similarity).  
 CC -!- TISSUE SPECIFICITY: Expressed predominantly in skeletal  
 CC muscle.  
 CC -!- SIMILARITY: BELONGS TO THE L56E FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AL553126; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; AL553134; -; NOT\_ANNOTATED\_CDS.

DR EMBL; AC026817; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; AK027808; BAB55384.1; ALT\_TERM.  
DR Genew; HGNC:16468; LACTB.  
DR InterPro; IPR001466; Beta\_lactamase.  
DR Pfam; PF00144; beta-lactamase; 1.  
KW Ribosomal protein; Mitochondrion.  
SQ SEQUENCE 547 AA; 60693 MW; C4BDE6B8EF39168A CRC64;

Query Match 68.8%; Score 33; DB 1; Length 547;  
Best Local Similarity 66.7%; Pred. No. 74;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VESYVPLFP 9  
|: ||| ||  
Db 189 VQHYVPEFP 197

Search completed: August 24, 2004, 15:43:45  
Job time : 8.83582 secs